APTD-0777

PB 204 148

COMPUTER MAPPING
as an aid in
AIR POLLUTION STUDIES

VOLUME TWO: INDIVIDUAL REPORTS

REPORT B
ST. LOUIS REGION STUDY -- PART II

by
Steven Fosburg

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
Springfield, Va. 22151

H1

54

BIBLIOGRAPHIC DATA SHEET	1. Report No. APTD-0777	2.	3. Recipient's Accession No.
4. Title and Subtitle Computer Mapping as an aid in Air Pollution Studies Volume Two: Individual Reports - Report B - St. Louis Region Study - - Part II			5. Report Date April 1970 6.
7. Author(s) Steven Fosburg			8. Performing Organization Rept. No.
9. Performing Organization Name and Address Graduate School of Design Harvard University Cambridge, Massachusetts 02138			10. Project/Task/Work Unit No. 11. Contract/Grant No. 68A-2405D
12. Sponsoring Organization Name and Address National Air Pollution Control Administration U. S. Public Health Service U. S. Department of Health, Education and Welfare Washington, D. C. 20201			13. Type of Report & Period Covered 14.

15. Supplementary Notes DISCLAIMER - This report was furnished to the Office of Air Programs by Graduate School of Deisgn, Harvard University, Cambridge, Massachusetts 02138 in fulfillment of Grant 68A-2405D

16. Abstracts

A report of an individual case study of the St. Louis region is presented. The report is divided into three sections; data analysis considerations, comparison of air quality and socio-economic data, and diurnal variations in air quality. The appendices include: a discussion of air quality surfaces; a discussion of socio-economic surfaces; and a presentation of computer graphics.

17. Key Words and Document Analysis. 17a. Descriptors

Air pollution
Computer programs
Mapping
Data processing
Socio-economic status
Sulfur dioxide
Particles

17b. Identifiers/Open-Ended Terms

St. Louis
Particulates

17c. COSATI Field/Group 13B

18. Availability Statement	Unlimited	19. Security Class (This Report) UNCLASSIFIED	21. No. of Pages 60
		20. Security Class (This Page UNCLASSIFIED	22. Price

FORM NTIS-35 (10-70)

USCOMM-DC 40329-P71

LABORATORY FOR COMPUTER GRAPHICS
AND SPATIAL ANALYSIS

Graduate School of Design
Harvard University
Cambridge, Massachusetts

This report was funded by:

Demonstration Grant 68A-2405D
National Air Pollution Control Administration
U. S. Public Health Service
U. S. Department of Health, Education, and Welfare

John C. Goodrich, Editor
Gail Howrigan, Associate Editor

Rosalie Alexander
Carol Chase
Candace Howes
Whitmore John
Linda Peacock

April 1970

COMPUTER MAPPING

as an aid in

AIR POLLUTION STUDIES

Personnel

Principal Investigators

Howard T. Fisher	Project Director (Associate Director of The Laboratory for Computer Graphics and Spatial Analysis)
John C. Goodrich	Associate Project Director
William Warntz	(Director of The Laboratory for Computer Graphics and Spatial Analysis)
Allan H. Schmidt	(Associate Director of The Laboratory for Computer Graphics and Spatial Analysis)

Case Study Investigators

Norma Bogen	St. Louis Project
Paul Jack Dangermond	California and Puget Sound Studies
Steven Fosburg	Statistical Investigation
Jean Maurice Granger	Montreal Project
Donald Shepard	St. Louis Project/Program Investigation
Elia Werczberger	Kansas City Project

Programming Support and Development

Chang-I Hua
Lannon Leiman
Kathleen Reine

Research Assistants

Whitmore John
Johnny Kirkland
Helen Mansfield
Adrian Thomas

Ellen Bussard
Robert Cartwright
Katherine Mickey
Warnie Webster

Administrative and Secretarial Support

Malcolm Fitzpatrick
Gail Howrigan
Linda Peacock
Laurence Yont

Regina McDonald
Donna L. Moorhead
Claudia S. Rosen
Judith Strong

PREFACE

From July 1, 1968 through December 31, 1969, the Laboratory for Computer Graphics and Spatial Analysis conducted a demonstration of several conceptual, methodological, and technical approaches to problems in the field of air pollution, employing computer-produced maps and allied displays. This work was done under a grant (No. 68A-2405D) from the Division of Control Agency Development, National Air Pollution Control Administration (NAPCA), Public Health Service, U. S. Department of Health, Education, and Welfare.

The results of this study, taken in its entirety, should make an important contribution to the work of Federal, state, and local air pollution agencies, especially in their efforts to determine the best control strategies for use in air quality control regions. The computer is an ideal tool for relating and combining sets of data, and for considering alternatives. For instance, graphic displays can contribute significantly to the evaluation of both short-term and long-term control strategies.

In demonstrating "the use of computer graphics as a means for more clearly portraying data related to the sources, distribution, and community effects of air pollution," the primary emphasis was on case studies for specific areas. The secondary emphasis, to support and complement the case studies, included:

(a) "the development of instructional manuals covering the procedures required for the preparation and interpretation of computer maps of various kinds, as related to air pollution,"

(b) "the use of certain newly developed techniques for describing, analyzing, and portraying quantitative surfaces and movements relative to such surfaces, as related to air pollution," and

(c) "an appraisal of alternative equipment, techniques, and computer programs presently or potentially available for the production of computer graphic displays pertinent to air pollution."

The case study regions were selected in consultation with NAPCA. Considerations included:

(a) sufficient air quality, emission, and related socio-economic and meteorological data--spatially distributed over the region--to enhance the likelihood of adequately demonstrating the mapping techniques;

(b) representative problems in data handling and mapping so that the techniques demonstrated would have applicability to as many other areas as possible;

(c) special problems in data handling and mapping so that as many techniques as possible could be adapted, developed, and tested;

(d) areas of which NAPCA and the individual investigators had special knowledge and could therefore utilize relevant research to enhance the study;

(e) areas where an immediate interest in mapping had been expressed for on-going Federal, state, and local air pollution activities.

Other research was undertaken to support and complement the case studies, and, in particular, to provide:

(a) the widest possible distribution of the mapping programs and accompanying explanation of the techniques through

i) copies of the Laboratory's SYMAP Computer mapping program,
ii) an introductory correspondence course in the SYMAP program,
iii) general program description and supplementary explanation for the existing user-manuals;

(b) a general review of present and potential hardware, output devices, and programs applicable to computer graphic displays of air pollution and related data;

(c) means for predicting, projecting, and correlating air pollution and related data, utilizing the latest

knowledge in planning, geography, and spatial analysis in general;

(d) dissemination of information on approaches and techniques to Federal, state, and local air pollution agencies, especially for use in the air quality control regions.

The final report is divided into two parts: Volume One: Summary, and Volume Two: Individual Reports--the detailed reports on the case study areas and the technical reports related to other research undertaken. The pertinent graphic displays are included with the appropriate individual report and a complete set is provided with the Summary Volume.

Introductory sections describing the air pollution and related data used are provided in the major case study reports for St. Louis and Kansas City; these should be referred to by anyone unfamiliar with the types of data and problems existing in this field. Detailed discussions of special data handling and mapping techniques are presented either as appendices to case study reports, or in separate technical reports.

The Summary Volume includes a discussion of all the areas of research undertaken and contains the complete table of contents for the reports of Volume Two. This should provide a useful reference of the special problems and technical discussions of one report which may be of interest to the reader of another detailed report.

The St. Louis Project was the major case study region examined under the demonstration grant. Consequently, the results are presented in three separate reports:

--Part I deals with the general mapping techniques;

--Part II with the statistical analyses undertaken; and

--Part III with a study of power plant load shifting.

The St. Louis Region Study--Part II--is presented here as Report B of Volume Two. It is divided into three sections--data analysis considerations, comparison of air quality and socio-economic data, and diurnal variations in air quality. An introduction discussing the area and the available data was included in Part I. Two technical appendices are included in addition to the appendix of computer graphics:

i) discussion of air quality surfaces;

ii) discussion of socio-economic surfaces.

The data used in the study was provided by the Division of Control Agency Development through the efforts of Robert Stephenson, Project Officer. The assistance of Jerry Mersch, NAPCA, the Department of Public Safety, Division of Air Pollution Control, City of St. Louis, and the Division of Air Pollution, St. Louis County Health Department, is also greatly appreciated.

This study was under the direction of Steven Fosburg, with the assistance of Norma Bogen and Lannon Leiman.

The report was prepared by Steven Fosburg.

John C. Goodrich, Editor
Cambridge, Massachusetts
April 1970

REPORT B
ST. LOUIS REGION STUDY -- PART II

Table of Contents

PREFACE

1. DATA ANALYSIS CONSIDERATIONS

   1.1 Generating Surfaces

   1.2 Problems of Missing Data

   1.3 Averaging of Data

   1.4 Inconsistent Data Bases

   1.5 Index of Air Pollution

2. COMPARISON OF AIR QUALITY AND SOCIO-ECONOMIC DATA

   2.1 Correlation with Sulfur Dioxide

   2.2 Regression Analysis - Sulfur Dioxide

   2.3 Regression Analysis - Suspended Particulates

3. DIURNAL VARIATIONS IN AIR QUALITY

REFERENCES

APPENDIX 1: COMPUTER GRAPHICS

APPENDIX 2: DISCUSSION OF AIR QUALITY SURFACES

APPENDIX 3: DISCUSSION OF SOCIO-ECONOMIC SURFACES

List of Illustrations

APPENDIX 1: COMPUTER GRAPHICS

Figure

B-1 Surface of 2-hour Sulfur Dioxide Readings, December 20, 1964--4, 6, and 8 a.m.; Interstate Study Area

B-2 Surface of 2-hour Sulfur Dioxide Readings, December 20, 1964--10, 12 a.m. and 2 p.m.; Interstate Study Area

B-3 2-hour SO_2 Surface in Oblique View, December 20, 1964--4 and 6 a.m.; Interstate Study Area

B-4 2-hour SO_2 Surface in Oblique View, December 20, 1964--8 and 10 a.m.; Interstate Study Area

B-5 2-hour SO_2 Surface in Oblique View, December 20, 1964--Noon and 2 p.m.; Interstate Study Area

B-6 Surface of 2-hour Sulfur Dioxide Readings, December 21, 1964--4, 6, and 8 a.m.; Interstate Study Area

B-7 Surface of 2-hour Sulfur Dioxide Readings, December 21, 1964--10, 12 a.m. and 2 p.m.; Interstate Study Area

B-8 St. Louis Interstate Air Pollution Study Area
- a. Regression Estimate of 24-hour SO_2 Readings
- b. 24-hour SO_2 Readings Corresponding to Above

B-9 1960 Census Population Density; Contour Map and Oblique View; SMSA

B-10 1960 Census Population Density--Oblique Views; SMSA
- a. 90^{o} Rotation
- b. 180^{o} rotation

B-11 1960 Census; Percent of Population under 10 Years of Age; SMSA

B-12 1960 Census; Percent of Population over 65 Years of Age; SMSA

B-13 1960 Census; Median School Years Completed; SMSA

B-14 1960 Census; SMSA
- a. Percent of Population Non-white
- b. Percent of Families with Incomes less than $3000/yr.

B-15 1960 Census; Percent of Families with Incomes less than $3000/yr.; Oblique View; SMSA

B-16 1960 Census; Density of Families with Incomes less than $3000/yr.; SMSA

B-17 1960 Census; SMSA
- a. Percent of Families with Incomes from $3000 to $7999/yr.
- b. Percent of Families with Incomes of $8000/yr. or more

B-18 1960 Census; Percent of Housing Units Deteriorating and/or Dilapidated; SMSA
- a. Contour View
- b. Oblique View

B-19 1960 Census; Percent of Housing Units with 1.01 of more Persons per Room; SMSA

B-20 1960 Census; Percent of Housing with 1 Unit/Structure; SMSA

B-21 1960 Census; SMSA
- a. Percent of Housing with 2 Units/Structure
- b. Percent of Housing with 3 or 4 Units/Structure

B-22 1960 Census; SMSA
- a. Percent of Housing with 5 to 9 Units/Structure
- b. Percent of Housing with 10 or more Units/Structure

B-23 1960 Census; Percent of Housing Built in 1939 or Earlier; SMSA

B-24 1960 Census; Percent of Housing Built 1940-1949; SMSA

B-25 1960 Census; Percent of Housing Built 1950-1960; SMSA

B-26 1960 Census; SMSA
- a. Density of Population over 65 Years of Age
- b. Oblique View of Same

B-27 1960 Census; SMSA
- a. Density of Population under 10 Years of Age
- b. Oblique View of Same

1. DATA ANALYSIS CONSIDERATIONS

Data handling and interpretation problems are inherent in most of the analyses done for this project. The manner in which these problems are solved--or, in some cases, bypassed--has a strong impact on the results produced. Frequently, either the available data are not very accurate or the necessary analysis technique is unknown or experimental. In such cases, careful attention must be paid to the interpretation of the results.

1.1 GENERATING SURFACES

For the two-hour sulfur dioxide data measured for St. Louis there were only eleven sampling stations and these were clustered near the central city.[1] From these eleven points we would like to determine how the air pollution behaves across the entire city and its environs. If we can think of a surface of air pollution in the same terms we use for physical terrain, then we are looking for the location and height of "mountains"--those areas having the highest concentration of air pollution. Subsequently, we will be concerned with who lives at the location of these mountains and how much time they spend there. This will tell us who is being exposed to pollution, and to what levels they are being exposed.

To determine the surface from these eleven points, we use a particular interpolation technique built into the SYMAP program.[2] The more points there are in an area, the more accurate the interpolation will be. The

interpolation is better between points than outside the area containing the data points. If the points are not equally spaced, different portions of the map will be accurate to different degrees. In no case were the pollution data measured on an equally spaced grid.

Since the concentration of observations is generally the greatest in the central city, the accuracy is usually best there. Outside the central core of observations, the extrapolated surface often becomes quite inaccurate. The extrapolated surface is an average of the nearest points; often this graphically produces areas of misleadingly high air pollution outside of the region we are examining--in areas where we would expect to find lower levels of pollution.

A great deal can be learned about the behavior of air pollution within the area containing the sampling stations and adjacent areas. If greater detail is desired, several options are available. We can (i) set up more sampling stations, or (ii) assume that air pollution behaves according to some known or hypothesized model. Available data can be fitted to such a model to estimate levels of air pollution for the region where the model is applicable. Such techniques as diffusion modeling, however, are quite complex and still too inaccurate for widespread use. Purchasing and operating more sampling stations is more immediately workable but expensive.

One solution would be to take infrequent observations around the periphery of the central city for a couple of years with mobile samples to estimate the background level for each season, and to project the background values

using straightforward statistical techniques. Background figures for each city are needed; it seems clear from what little national data we have that the background levels vary for different parts of the country.

1.2 PROBLEMS OF MISSING DATA

Frequently, missing observations force us to deal with long-term averages, although in some instances they can be replaced by estimates of the data. For example, the SO_2 levels for each of the eleven sampling stations in St. Louis was averaged over all the observations at a particular station. Since the number of observations differed from station to station, the accuracy--standard deviation--of each of these averaged values is different. Fortunately, there seemed to be no pattern to the missing data. If weekday data is substituted for missing weekend data the resulting overall average will be higher than a true weekly average. Including stations like this in an analysis of the air pollution surface can give very misleading results.

Diurnal data for SO_2 (see Figures B-1 to B-7) were measured at only eleven stations every two hours; consequently, the interpolated area around any missing value could be dangerously misleading. If the missing value were on what was actually a peak, the peak would not appear; if it were in a valley, the elevation of the bottom of the valley would be over-estimated by the interpolation process. Despite the anomalies caused by the missing data, it was clear that the pollution surface shifted considerably during the day.

1.3 AVERAGING OF DATA

Either standard arithmetic means or geometric means are used when the data are to be averaged. Since it is generally acknowledged that concentrations of air pollutants are approximately lognormally distributed for all averaging times, the geometric mean is often used.[3] In our study we used both standard and geometric means.

1.4 INCONSISTENT DATA BASES

The data for different pollutants were collected from incommensurate sets of sampling stations. Since the sulfur dioxide and particulate readings were made at different sets of points, comparisons of these two pollutants are not reliable considering how few stations were used.

In Section 2 two incommensurate data sets are discussed. The air pollution data was measured at the various sampling stations and was related to the socio-economic data compiled by Census tracts. There are many more Census tracts than there are sampling stations, and the data points do not coincide. To compare the different sets of data, we interpolate, using the more accurate Census tract data. At each point in the study area at which a sampling station is located, we compute the estimated value of the Census tract data. We now have two commensurate sets of data -- the actual air pollution data and the estimated values of the socio-economic data -- which can be compared.[4]

1.5 INDEX OF AIR POLLUTION

Presently it appears that air quality is measured under the assumption that the effects of the pollutants can be considered separately. This is unrealistic because many of the pollutants coexist in the atmosphere and may act synergistically. The analysis of individual pollutants without regard for their combined effects should be considered a preliminary step to a pollution index. Once we have such an index,[5] we can reasonably decide upon levels which should not be tolerated and upon the most economical approach for bringing levels of air pollution within tolerable limits. For example, the pollution index might indicate that to obtain a certain reduction in the pollution index, one could either reduce sulfur dioxide by a certain amount, or nitric oxide by five times that amount. Without an index giving the relationships among the pollutants -- in terms of their effects upon human beings, plants, or materials -- the control procedure is apt to be much more subjective and arbitrary.

In Section 2 particular pollutants are related to various socio-economic parameters. Had a pollution index been available, it should have been used, rather than the individual pollutants.

2. COMPARISON OF AIR QUALITY AND SOCIO-ECONOMIC DATA

The relationships between air pollution and various socio-economic statistics -- pertaining to the population and housing in the St. Louis study area -- were analyzed to ascertain who and what are actually being affected by air pollution. Implicit in such a study is the assumption that meaningful control procedures can be carried out only when the tolerance levels and the possible kinds of long range effects are known.

The study was limited by a lack of data in usable form. Such important variables as land use were not analyzed in detail since the data existed in graphic rather than numeric form. Since the socio-economic variables from the 1960 U.S. Census were to be compared with data collected three, four, and five years later, we had to assume that the data did not drastically change during this period. Unfortunately, we do not have sufficient data to know how "stable" air pollution surfaces are over time -- relative to the surfaces of housing and population. One of the data sets used was the twenty-four hour dioxide readings, taken only during the winter months for two years. The average surface over the entire year might be quite different and lead to different conclusions.

Since there seemed to be no precedent for this analysis, it was necessary to examine many variables so as not to eliminate a variable that might later prove to be important. The variables that were chosen are

described in Appendix 3. Stepwise linear regression was used to find the best combination of variables. The one variable that is most highly correlated with the air quality data is found first. If two variables are required, a second variable is found which -- in combination with the first -- explains more variation in the data than any other combination. If three variables are needed, a third variable is found in a like manner, and so on, depending upon the degree of explanation required. The number of variables entered into the regression should be limited for statistical reasons such that there are many more data points than variables. For the analysis of sulfur dioxide there were 37 data points used, and 14 for the suspended particulates.

On occasion, a variable may appear to be completely explained by later variables entered into the equation; in this case, the explained variable should be removed so as to decrease the standard error of the estimating equation.

2.1 CORRELATION WITH SULFUR DIOXIDE

The variables referred to in Appendix 3 were compared with both the arithmetic and geometric means of the twenty-four hour sulfur dioxide readings discussed in Appendix 2, resulting in a twenty by twenty table of correlations. The variables most highly correlated with the geometric mean are also most highly correlated with the arithmetic mean.

The strongest correlations with the geometric mean were percent of non-White population (correlation .708),

percent income under $3000 (.768), and percent of deteriorating or delapidated housing (.740). None of these variables individually accounts for a sufficient amount of the variability in the air pollution data. Values around .9 would be accepted as adequate. However, variables which are individually inadequate can often explain the variability much better when combined.

2.2 REGRESSION ANALYSIS - SULFUR DIOXIDE

The results of the analysis of the twenty-four hour sulfur dioxide data were not significant. The variables were entered into the equation in the following order:

(i) Percent income under $3000 (Figure B-15)
(ii) Percent population non-White (Figure B-14a)
(iii) Percent housing built in 1939 or earlier (Figure B-23)
(iv) Percent housing which is deteriorated and delapidated (Figure B-18)
(v) Percent housing with 1.01 persons per room or more (Figure B-19).

The multiple R achieved with these five variables was only .7939. (Examine Figures B-8a and B-8b to visually compare the regression with the actual 24 hour SO_2 surface. Furthermore, it is interesting to see how the peaks in the maps referenced above compare with the SO_2 map.) It is also interesting to note that these five variables are all associated with poverty.

2.3 REGRESSION ANALYSIS - SUSPENDED PARTICULATES

Suspended particulates data were measured from May 20, 1963 to April 21, 1964. (See Figures A-13a and A-13b.) The number of readings per station varied from a high of 207 to a low of 57, producing disparities in the accuracy with which the surface of suspended particulates is estimated over this period. Ideally, we would like to have at least a complete year's records from all the stations for a city.

The correlations between the socio-economic variables and the arithmetic means of the suspended particulate data differed from the results found with the sulphur dioxide data. There were only six individual correlations with absolute values greater than .5, including percentage of housing built from 1950 to 1960 (.594), median school years completed (-.633), percentage of housing built in 1939 or earlier (.572), and percentage of housing with 1.01 or more persons per room (.570). One must be careful to note that some of the correlations above are negative. This means that the variables are inversely correlated. For example, the surface for median school years completed -- correlation of -.633 -- is highest in parts of the suburbs where suspended particulate concentration is relatively low.

Although individually the variables do not explain the data very well, there is a considerable degree of explanation when the variables are in combination.

The variables were entered in the following order:

(i) Median school years completed (Figure B-13)

(ii) Percent of population with incomes of $8000 or more (Figure B-17b)

(iii) Percent of housing built, 1950-1960 (Figure B-25)

(iv) Density of population sixty-five years old and older (Figure B-12).

The multiple R achieved here is quite high -- .9106. The standard error of the estimate is 25.98.

These results should not be construed to indicate that the four variables above are unique in their explanatory power, especially since the variables are redundant.[6] To demonstrate this fact, an alternate set of four variables was used to explain the data, entered in the following order:

(i) Density of the population aged sixty-five and older (Figure B-26)

(ii) Percentage of housing built in 1939 or before (Figure B-23)

(iii) Percentage of the population under ten years of age (Figure B-11)

(iv) Percentage of housing with two units in the structure (Figure B-26).

The contour maps of these variables are similar to the first set. The multiple R achieved was 0.9182 -- somewhat higher than before; the standard error of the estimate is 24.90 -- somewhat lower than before. The multiple R in this case is much higher than that found for the sulfur dioxide data. This is due both to better explanation and to the fact that there is less data here to explain, since there were so few sampling stations.

3. DIURNAL VARIATIONS IN AIR QUALITY

In studying air pollution, we are concerned not only with the average levels of pollution, but more importantly with the actual levels at any one time, since it is the actual levels that affect people. Furthermore, we are concerned with those areas where the levels are highest -- the locations of the peaks on the air pollution surface. If the major contribution to a high average level of sulfur dioxide comes from peaks in a relatively uninhabited area, we should be less concerned than if the major contribution to a lower average level of SO_2 comes from peaks in an area highly populated during the day or night.

Contour and oblique view maps were produced for a two day period from the two hour sequential sulfur dioxide data for St. Louis. These maps clearly showed how the levels of air pollution change during the day and what areas are most affected. (See Figures B-1 to B-7.) Maps such as these may some day be a normal part of every weather forecast. Preliminary investigation of the results indicated that the changes in the surface during the day were due not to winds and other such climatological factors, but mainly to differences in emissions. This would favor control legislation which would regulate the times at which certain emitters could operate at peak capacity.

REFERENCES

(1) The pollution data available for the St. Louis study are summarized briefly in Appendix 2. When problems of analysis are discussed in this section, reference is made primarily to this body of data. The applicable socio-economic data is summarized in Appendix 3. The St. Louis area and the available data are discussed in greater detail in Volume Two, Report A.

(2) See Volume Two, Report L for discussions of the interpolation function employed in SYMAP.

(3) See Appendix 4 of the report on the Kansas City Project, Volume Two, Report D.

(4) There are several alternative ways of producing commensurate sets of data; this was thought to be the most straightforward method.

(5) A problem of interest is the creation of a measure which more accurately embodies what is meant by the notion of air pollution. This measure should reflect the total effect of the separate air pollutants on property, people, or possibly both, in a fashion analogous to the temperature-humidity index.

A first approach to this problem would be to create an experimental environment in which the levels of several air pollutants could be controlled. At each level for various periods of time human observers would be asked for their reactions to the present level of air quality. The effects of air pollution on plants and/or materials could be taken at the same time. One or more air pollution indices could be derived by correlating all of these observations.

(6) For example, the region for which the percentage of housing units with one unit in the structure is high is the same as the region for which the percentage of housing with two units in the structure is low. If the contour map of one of these variables were printed with the symbolism interchanged -- high symbolism for the low values, and vice versa -- it would be hard to distinguish between the maps.

APPENDIX ONE

COMPUTER GRAPHICS

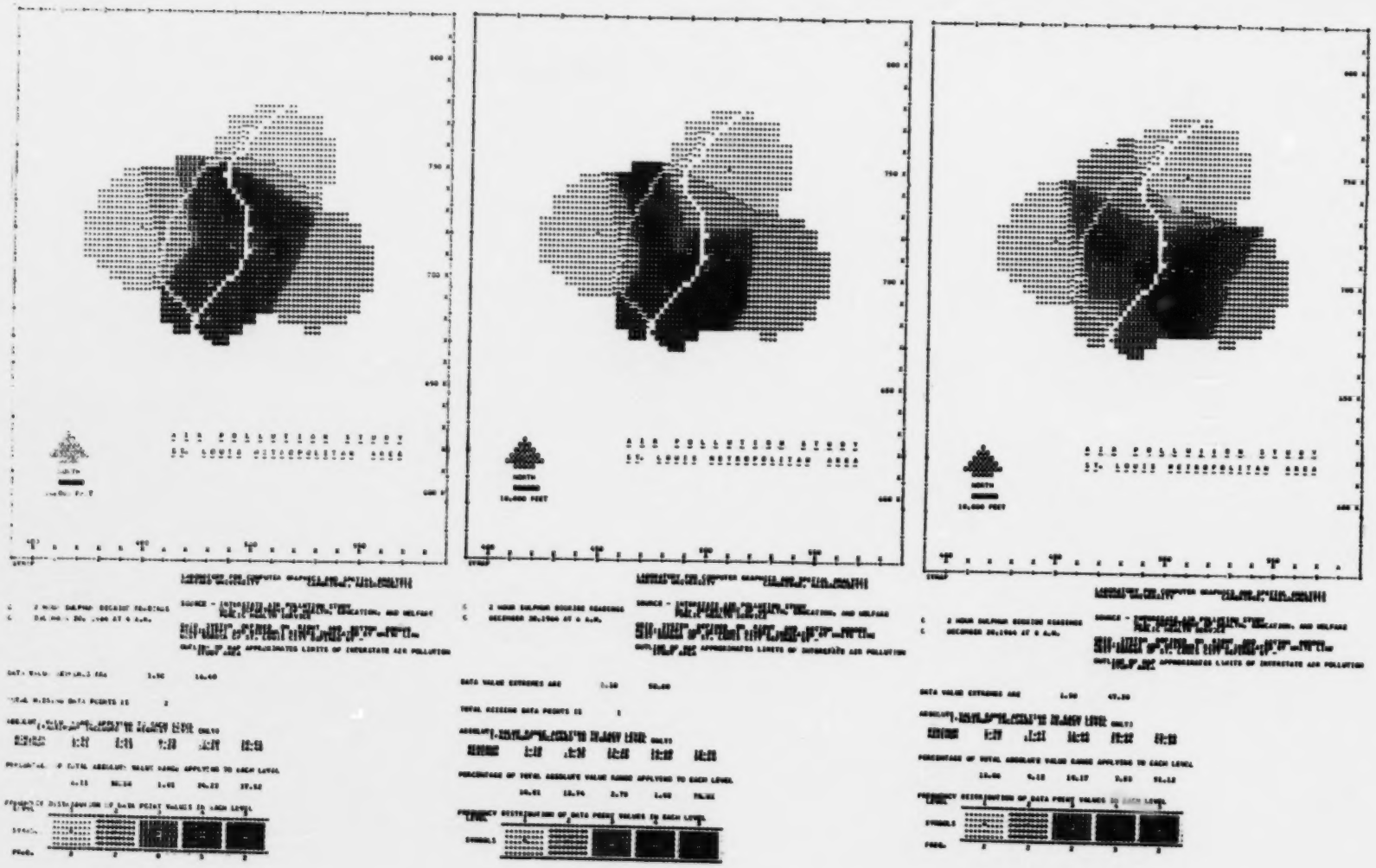


FIGURE B-1

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
Surface of 2-hour Sulfur Dioxide Readings
December 20, 1964 — 4, 6, and 8 a.m.

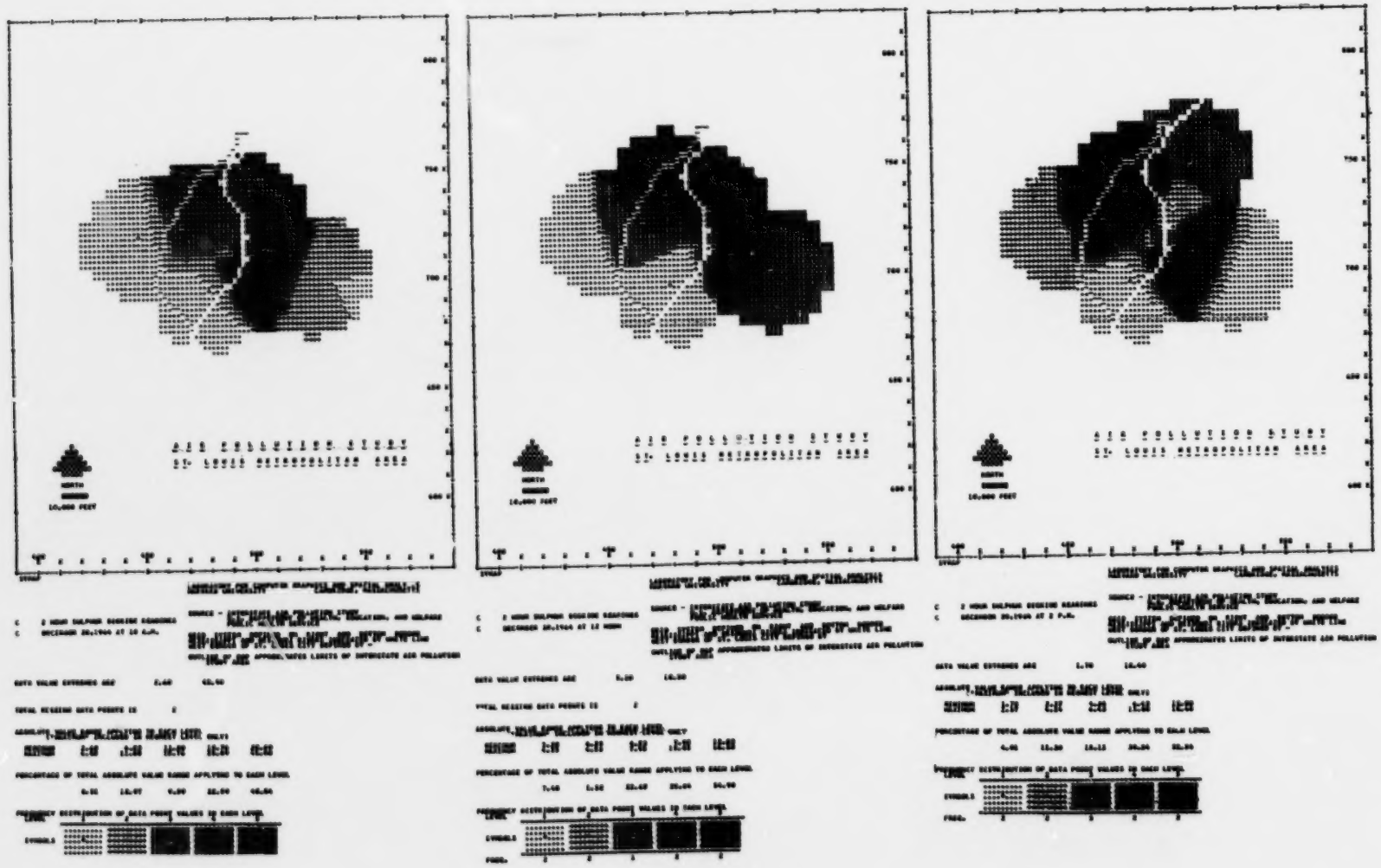


FIGURE B-2

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
Surface of 2-hour Sulfur Dioxide Readings
December 20, 1964 — 10, 12 a.m. and 2 p.m.

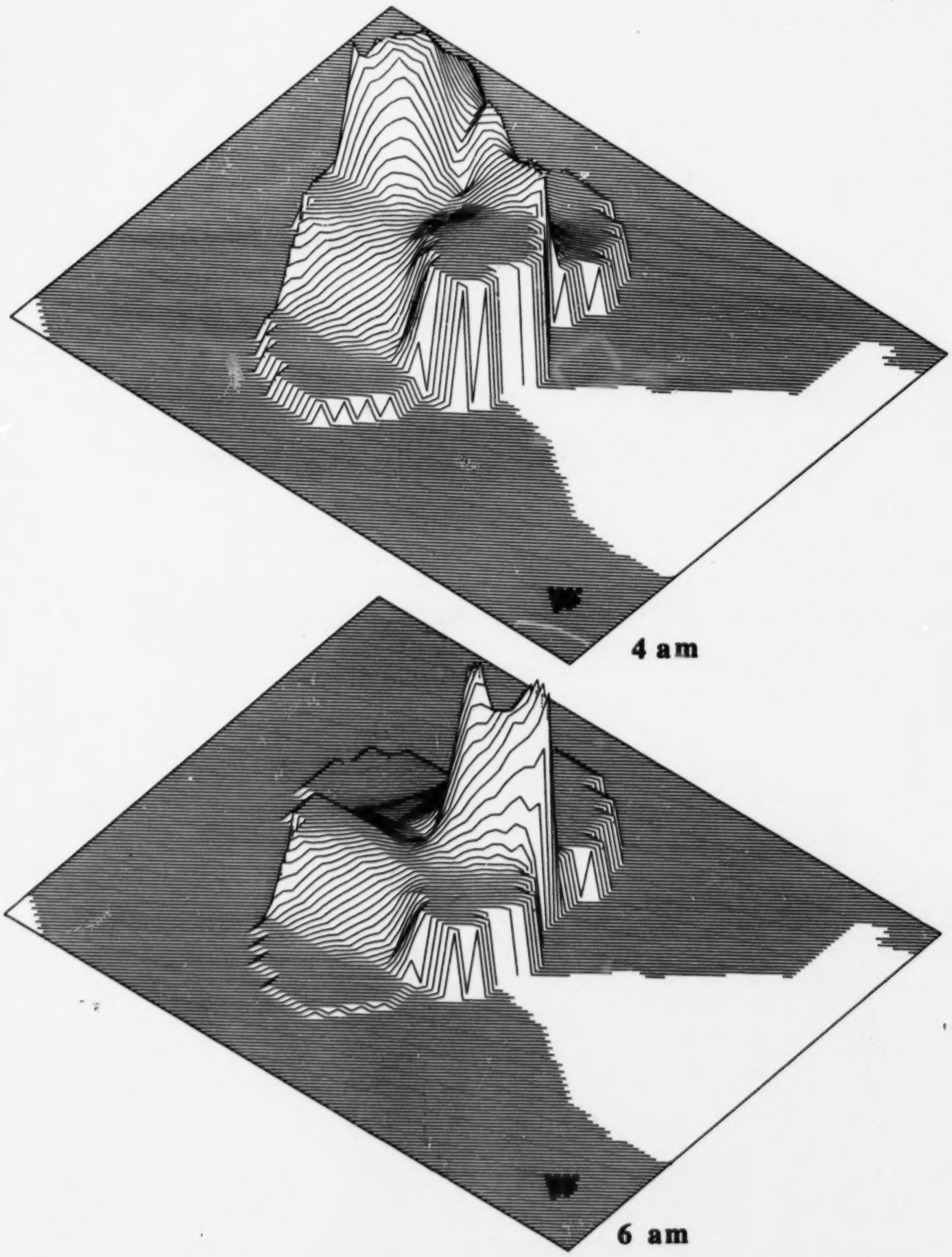


FIGURE B-3

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
2-hour SO_2 Surface in Oblique View
December 20, 1964 — 4 and 6 a.m.

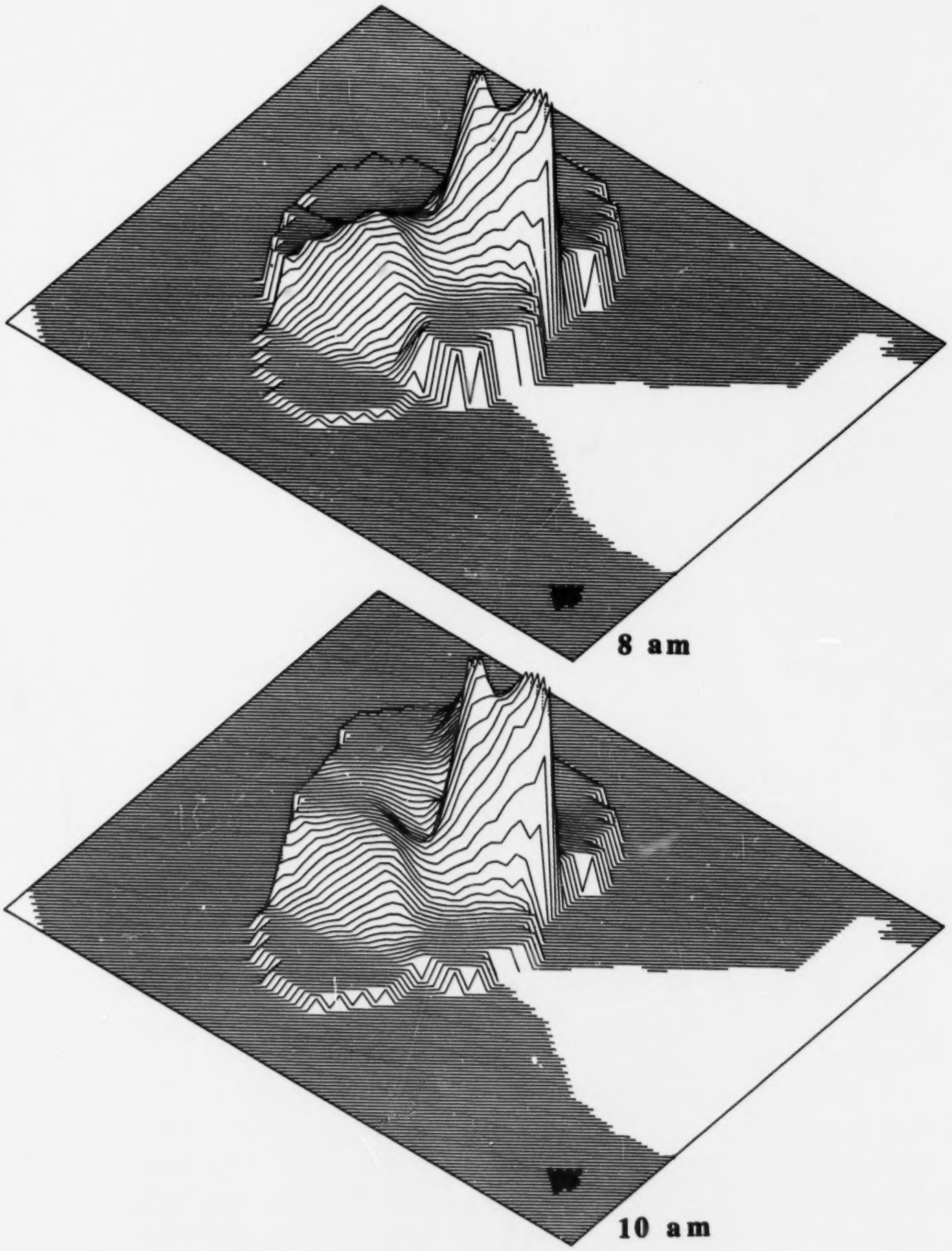


FIGURE B-4

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
2-hour SO_2 Surface in Oblique View
December 20, 1964 — 8 and 10 a.m.

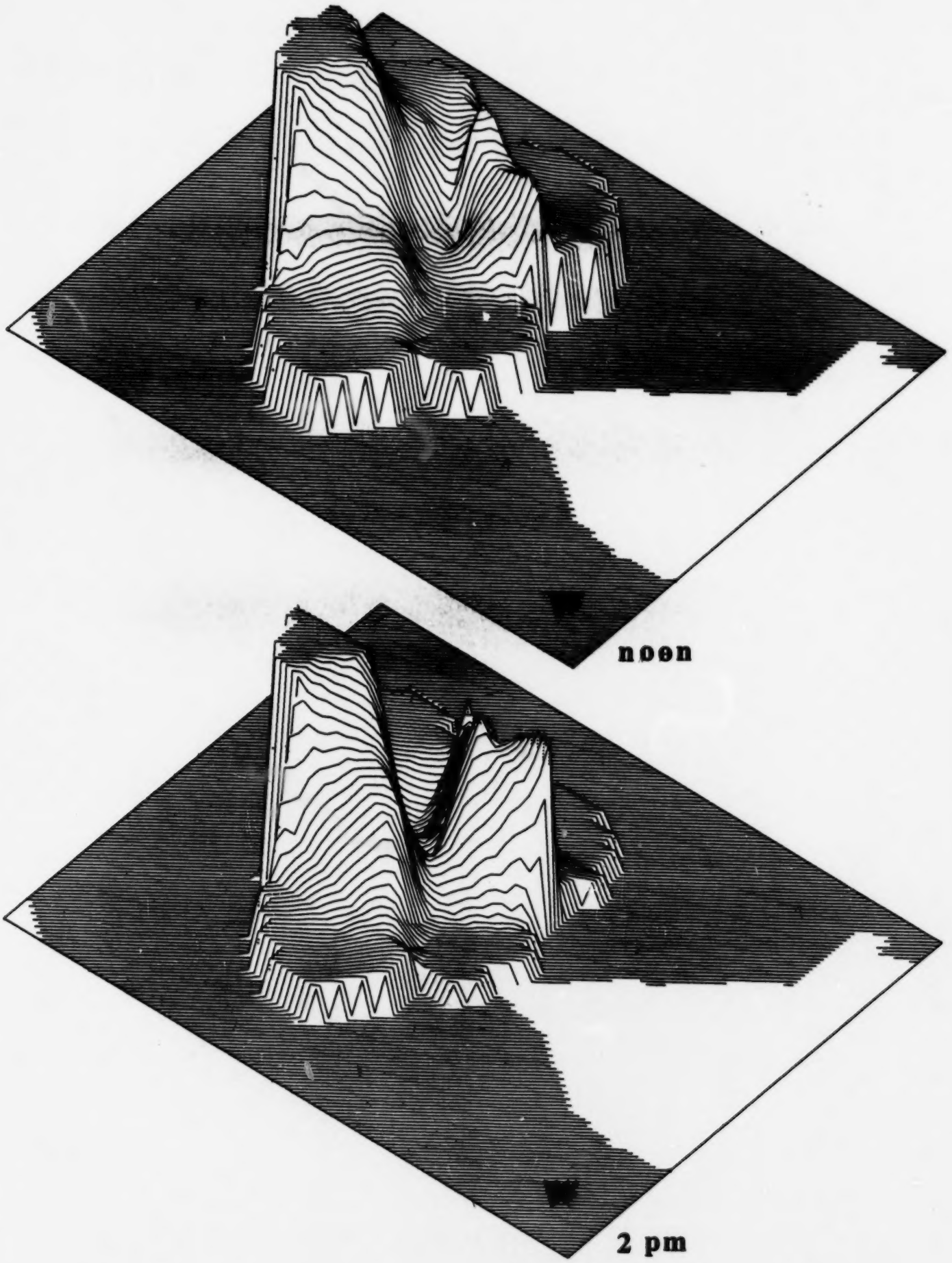


FIGURE B-5

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
2-hour SO_2 Surface in Oblique View
December 20, 1964 — noon and 2 p.m.

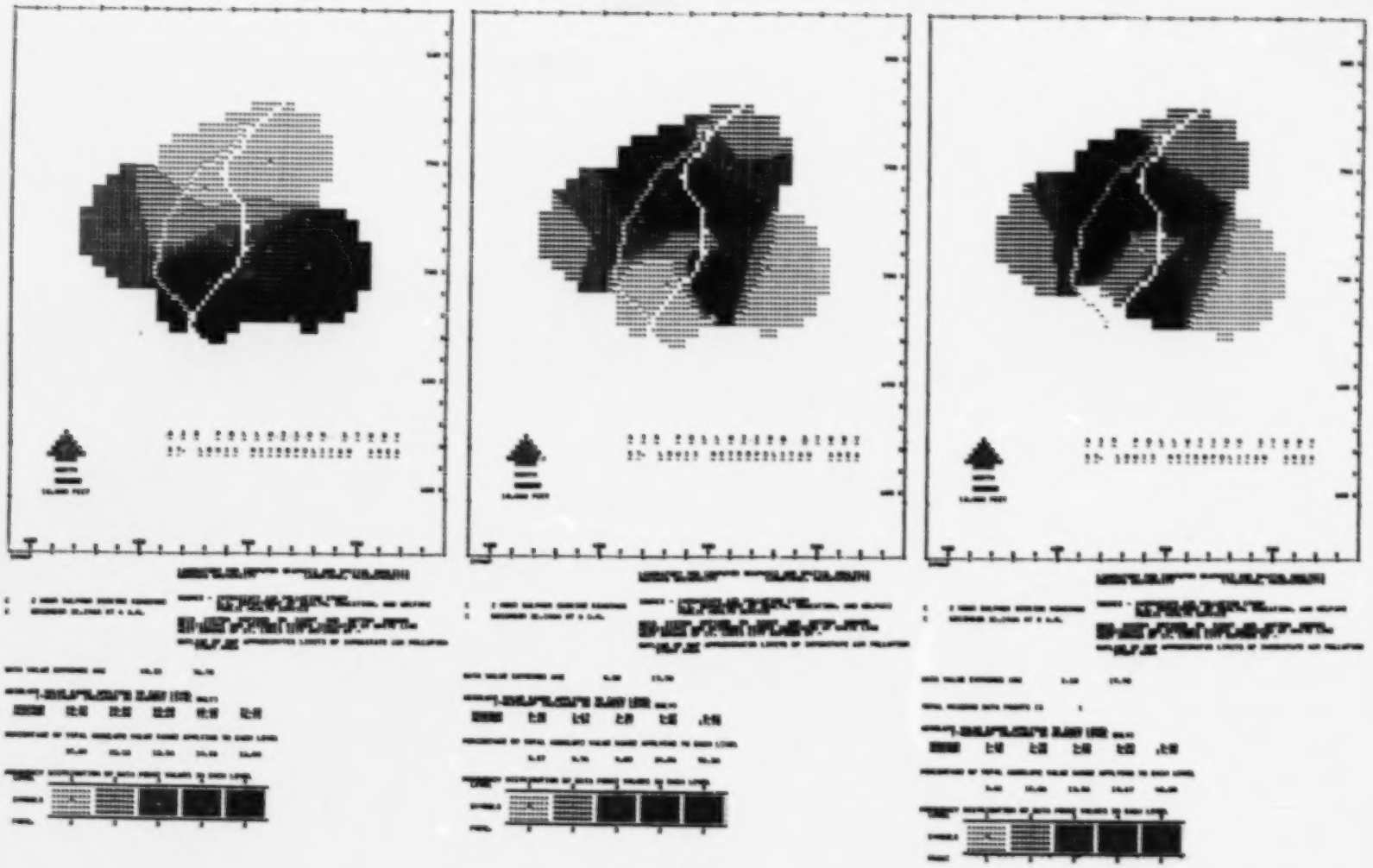


FIGURE B-6

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
Surface of 2-hour Sulfur Dioxide Readings
December 21, 1964 — 4, 6, and 8 a.m.

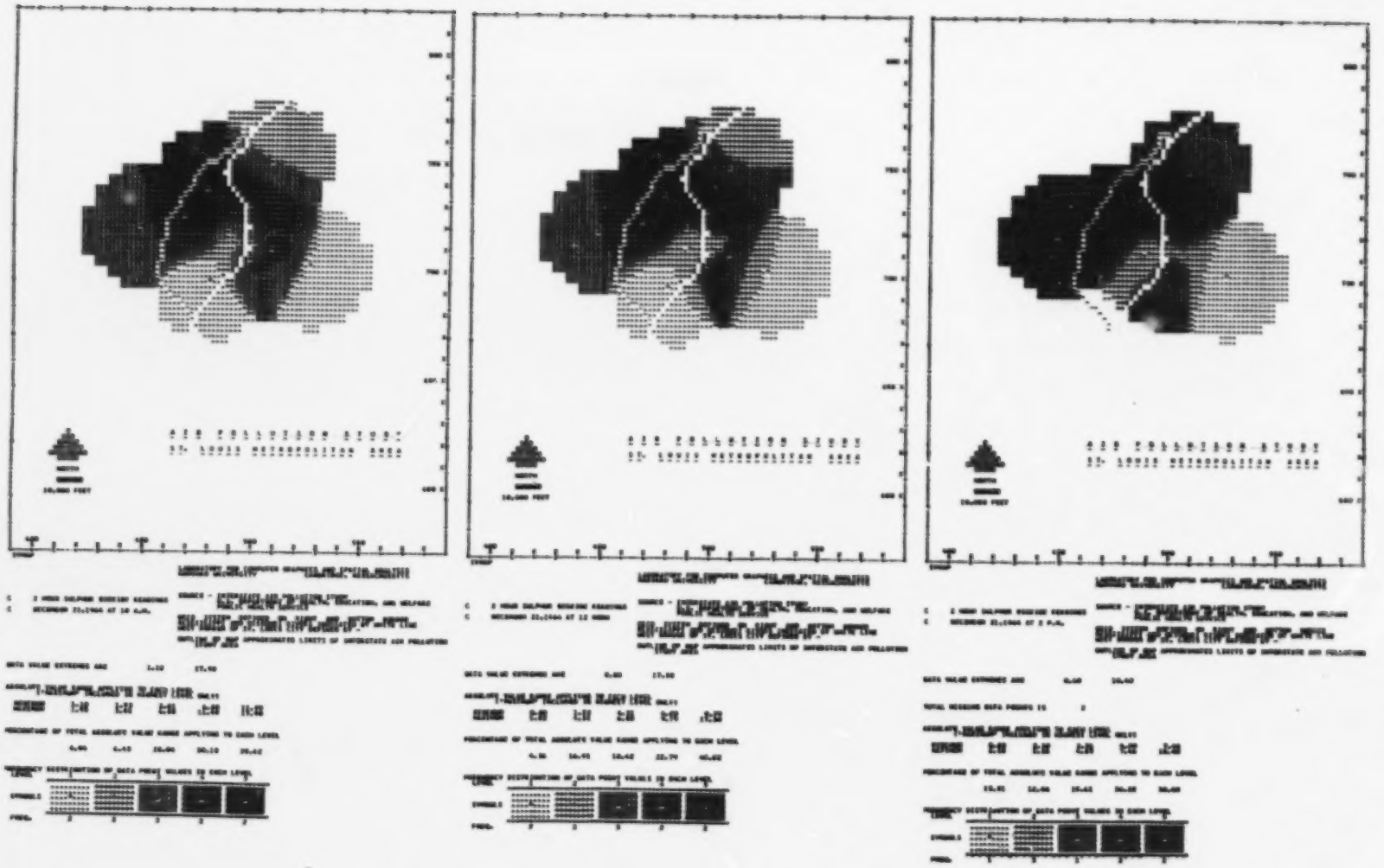
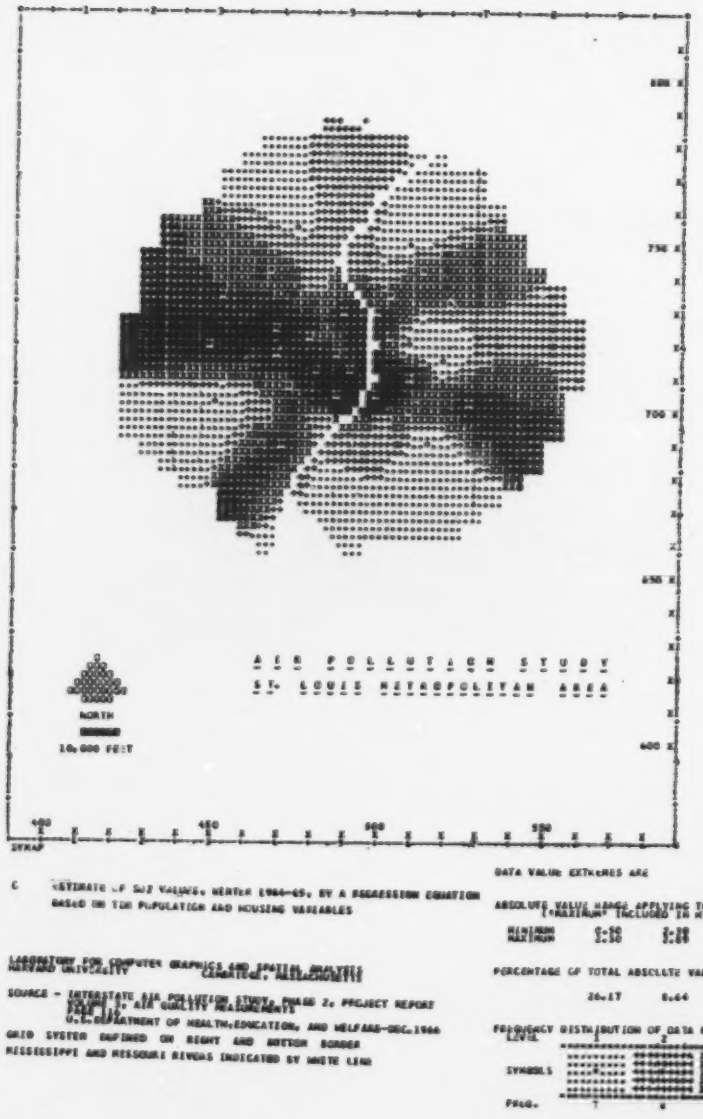


FIGURE B-7

ST. LOUIS REGION STUDY
St. Louis Interstate Air Pollution Study Area
Surface of 2-hour Sulfur Dioxide Readings
December 21, 1964 — 10, 12 a.m. and 2 p.m.


AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
NORTH
10,000 FEET
SYRAP
C ESTIMATE OF SO2 VALUES, WEATH 1964-65, BY A REGRESSION EQUATION BASED ON THE POPULATION AND HOUSING VARIABLES
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS HARVARD UNIVERSITY CAMBRIDGE, MASSACHUSETTS
SOURCE - INTERSTATE AIR POLLUTION STUDY, PHASE 2, PROJECT REPORT PUBLIC HEALTH SERVICE PUBLICATION NO. 999-AP-35 U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE-DEC. 1966
GRID SYSTEM DEFINED ON RIGHT AND BOTTOM BORDER
MISSISSIPPI AND MISSOURI RIVERS INDICATED BY WHITE LINE
DATA VALUE EXTREMES ARE 0.50 7.31
ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL (*MAXIMUM INCLUDED IN HIGHEST LEVEL ONLY)
MINIMUM 1:50 2:28 3:07 4:48 5:89 MAXIMUM
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
26.27 8.64 16.24 12.68 36.27
FREQUENCY DISTRIBUTION OF DATA POINT VALUES IN EACH LEVEL
LEVEL 1 2 3 4 5
SYMBOLS
FREQ. 7 4 7 6 7

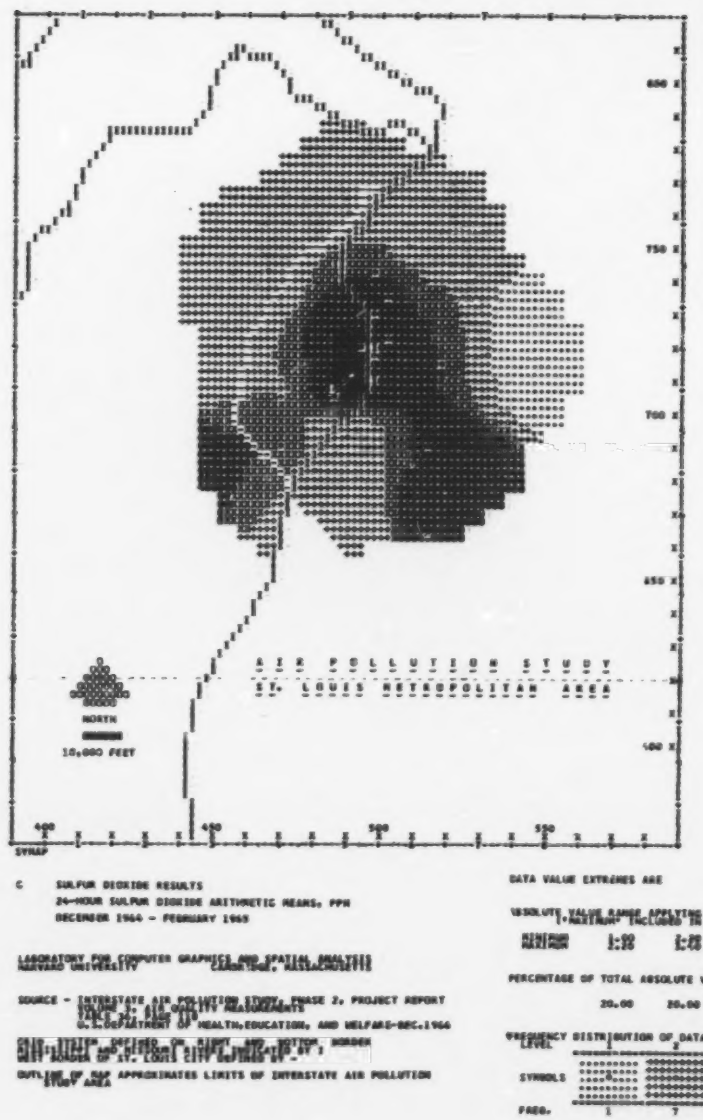

AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
NORTH
10,000 FEET
SYRAP
C SULFUR DIOXIDE RESULTS
24-HOUR SULFUR DIOXIDE ARITHMETIC MEANS, PPM
DECEMBER 1964 - FEBRUARY 1965
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS HARVARD UNIVERSITY CAMBRIDGE, MASSACHUSETTS
SOURCE - INTERSTATE AIR POLLUTION STUDY, PHASE 2, PROJECT REPORT PUBLIC HEALTH SERVICE PUBLICATION NO. 999-AP-35 U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE-DEC. 1966
GRID SYSTEM DEFINED ON RIGHT AND BOTTOM BORDER
WEST BORDER OF ST. LOUIS CITY COINCIDES WITH 90°
OUTLINE OF AQCR APPROXIMATES LIMITS OF INTERSTATE AIR POLLUTION STUDY AREA
DATA VALUE EXTREMES ARE 2.00 6.40
ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL (*MAXIMUM INCLUDED IN HIGHEST LEVEL ONLY)
MINIMUM 2:00 3:00 4:00 5:00 6:00 MAXIMUM
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
20.00 20.00 20.00 20.00 20.00
FREQUENCY DISTRIBUTION OF DATA POINT VALUES IN EACH LEVEL
LEVEL 1 2 3 4 5
SYMBOLS
FREQ. 1 7 3 6 2

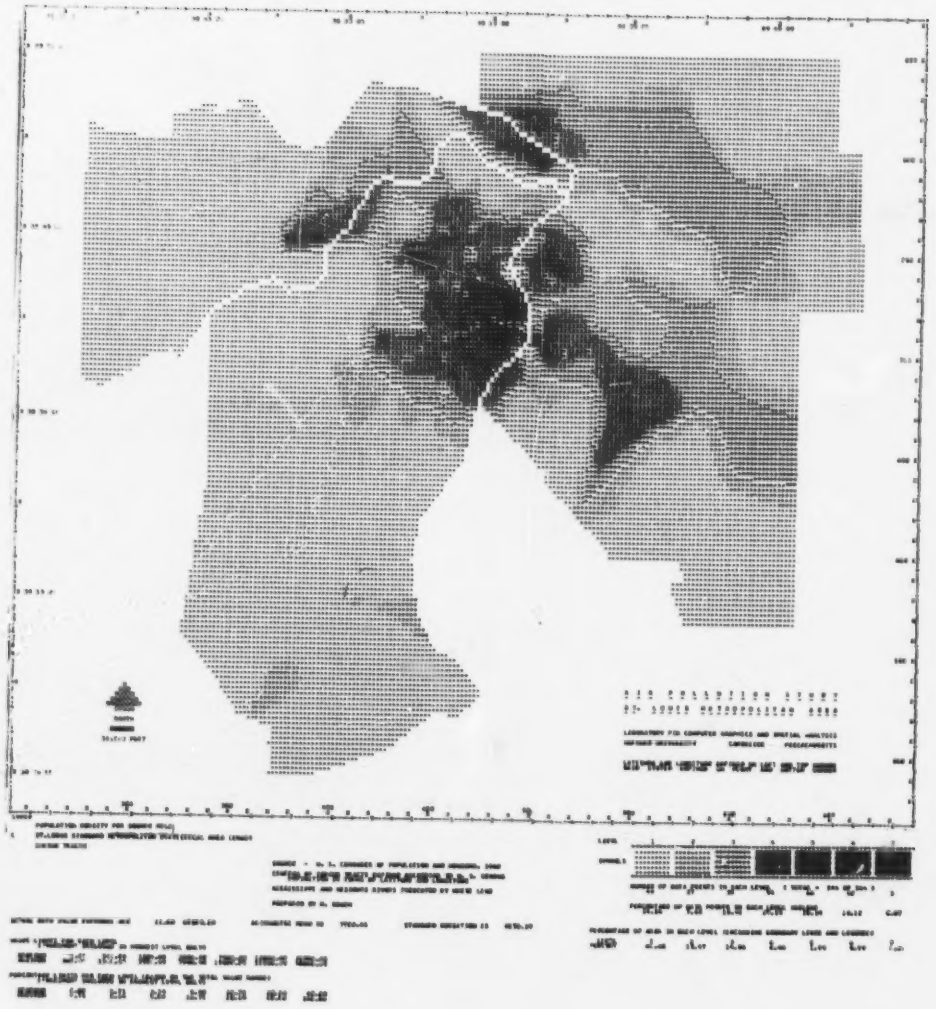
FIGURE B-8

ST. LOUIS REGION STUDY

St. Louis Interstate Air Pollution Study Area

a. Regression Estimate of 24-hour SO_2 Readings

b. 24-hour SO_2 Readings Corresponding to Above



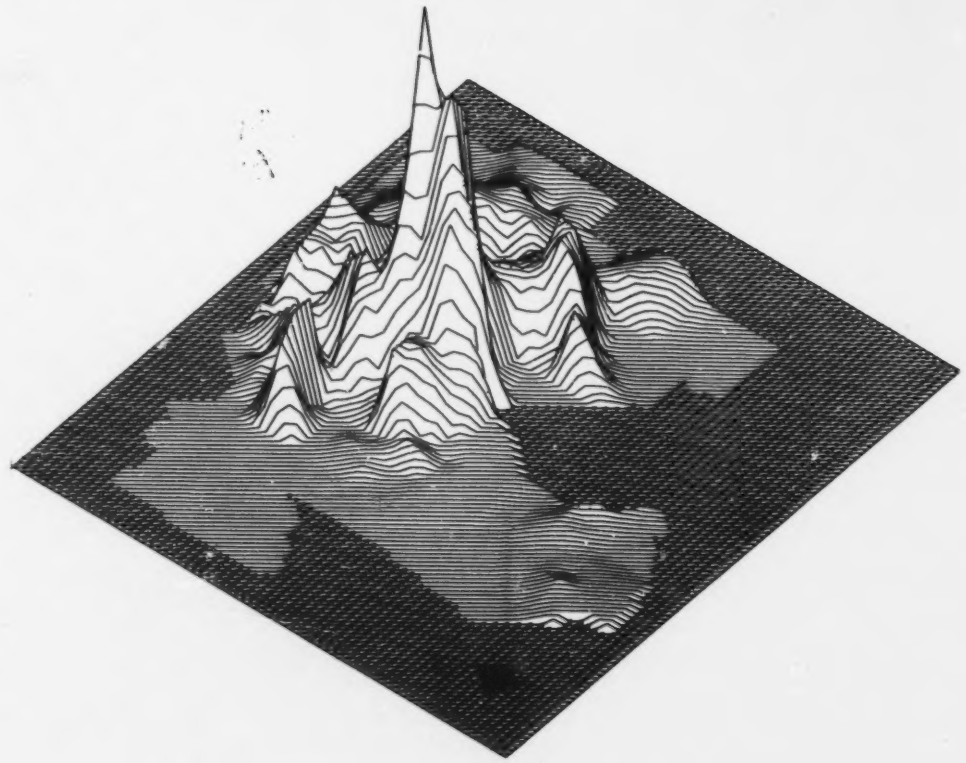


FIGURE B-9

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Population Density
Contour Map and Oblique View

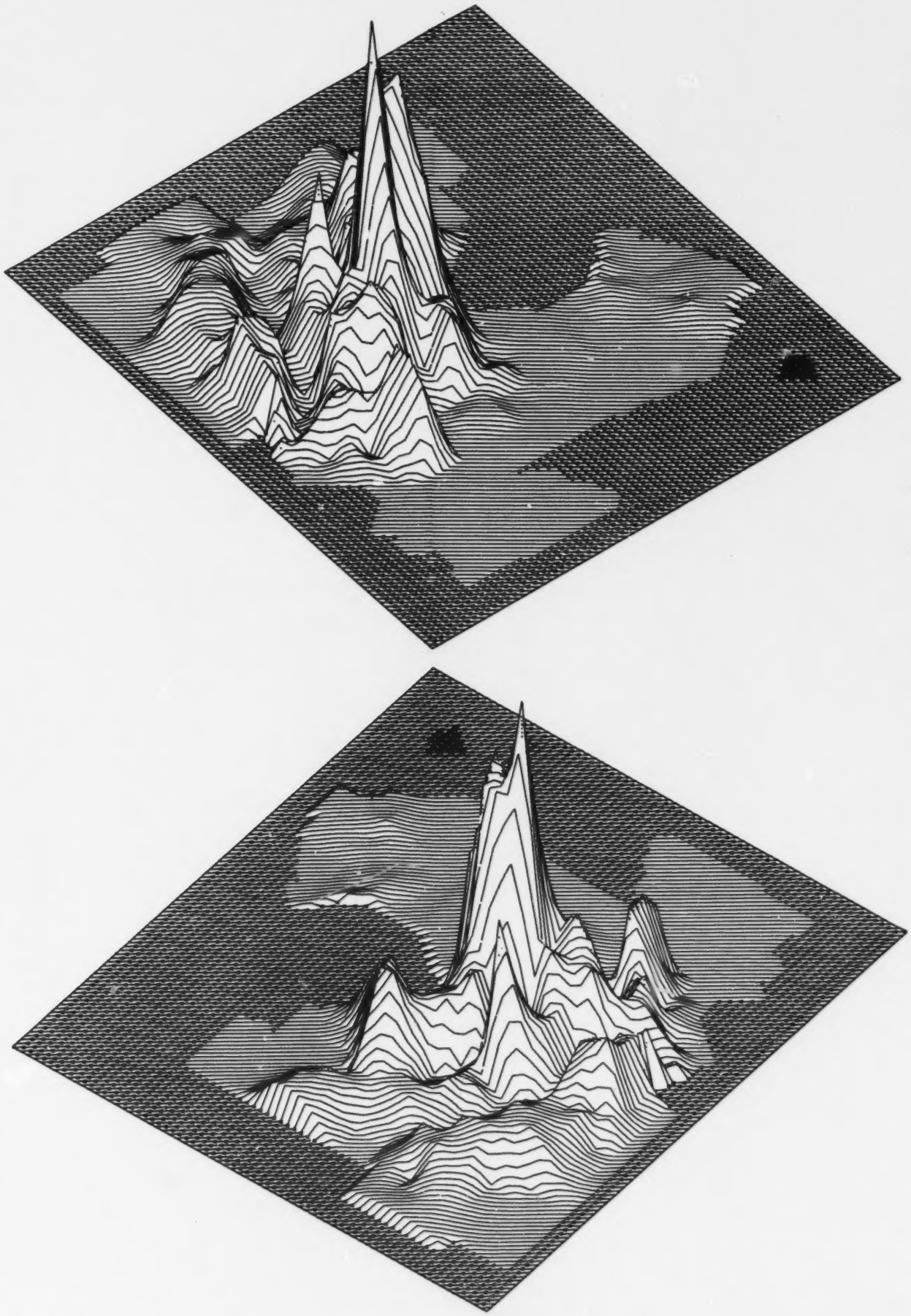


FIGURE B-10

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Population Density — Oblique Views
a. 90° Rotation b. 180° Rotation

AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS
HARVARD UNIVERSITY CAMBRIDGE MASSACHUSETTS
NORTH
10,000 FEET
PERCENT POPULATION UNDER TEN YEARS OF AGE
ST.LOUIS STANDARD METROPOLITAN STATISTICAL AREA (1960)
CENSUS TRACTS
SOURCE - U. S. CENSUSES OF POPULATION AND HOUSING, 1960
DATA VALUE EXTREMES ARE 0.03 39.73
PERCENTAGE OF TOTAL AREA
20.00 20.00 20.00 20.00 20.00
SYMBOLS

FIGURE B-11

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Population under 10 Years of Age

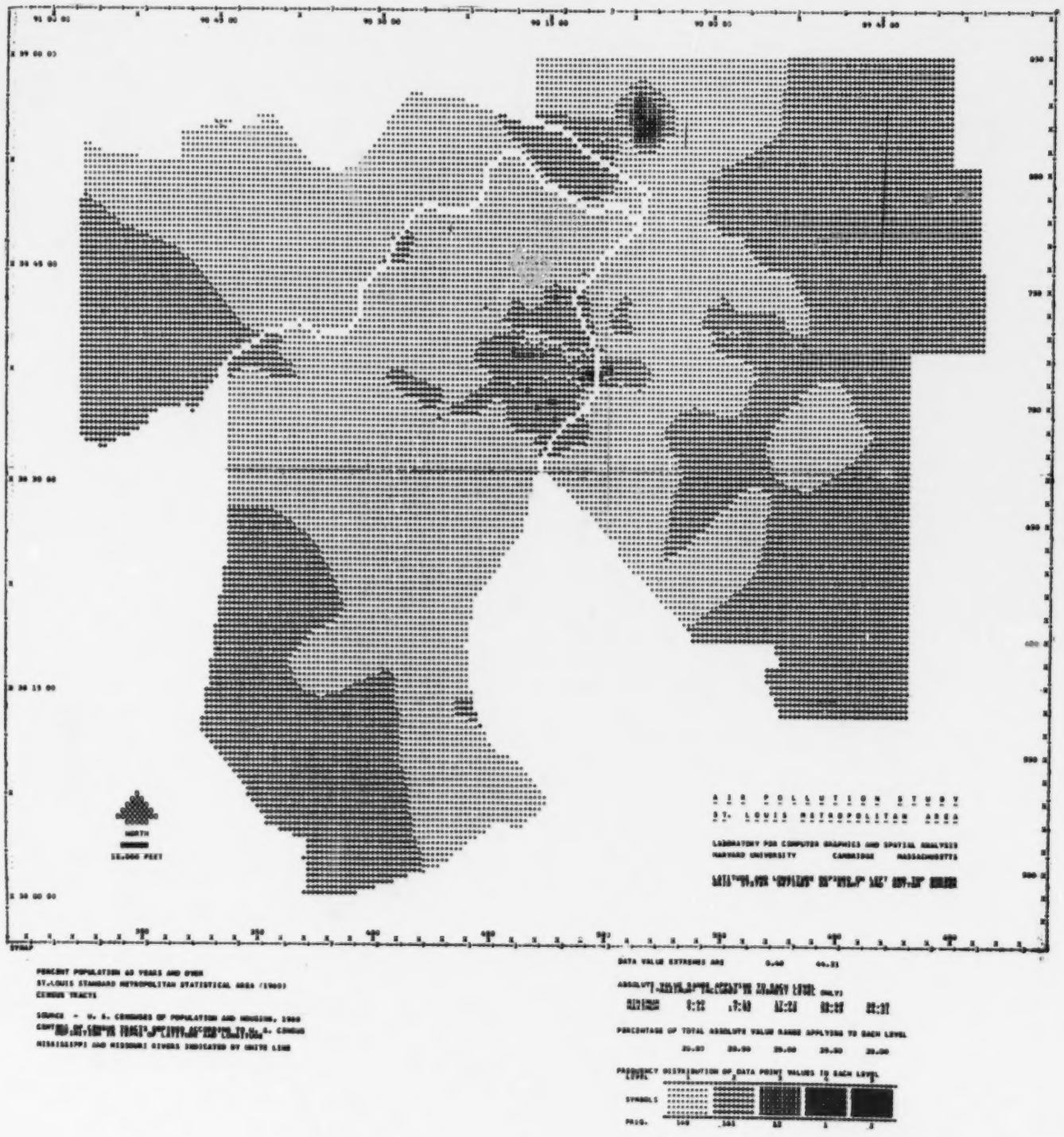

AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS
HARVARD UNIVERSITY CAMBRIDGE MASSACHUSETTS
NORTH
10,000 FEET
PERCENT POPULATION 65 YEARS AND OVER
ST.LOUIS STANDARD METROPOLITAN STATISTICAL AREA (1960)
CENSUS TRACTS
SOURCE - U. S. CENSUSES OF POPULATION AND HOUSING, 1960
MISSISSIPPI AND MISSOURI RIVERS INDICATED BY WHITE LINE
DATA VALUE EXTREMES ARE 0.40 44.31
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
20.00 20.00 20.00 20.00 20.00
SYMBOLS
FREQ.

FIGURE B-12

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Population over 65 Years of Age

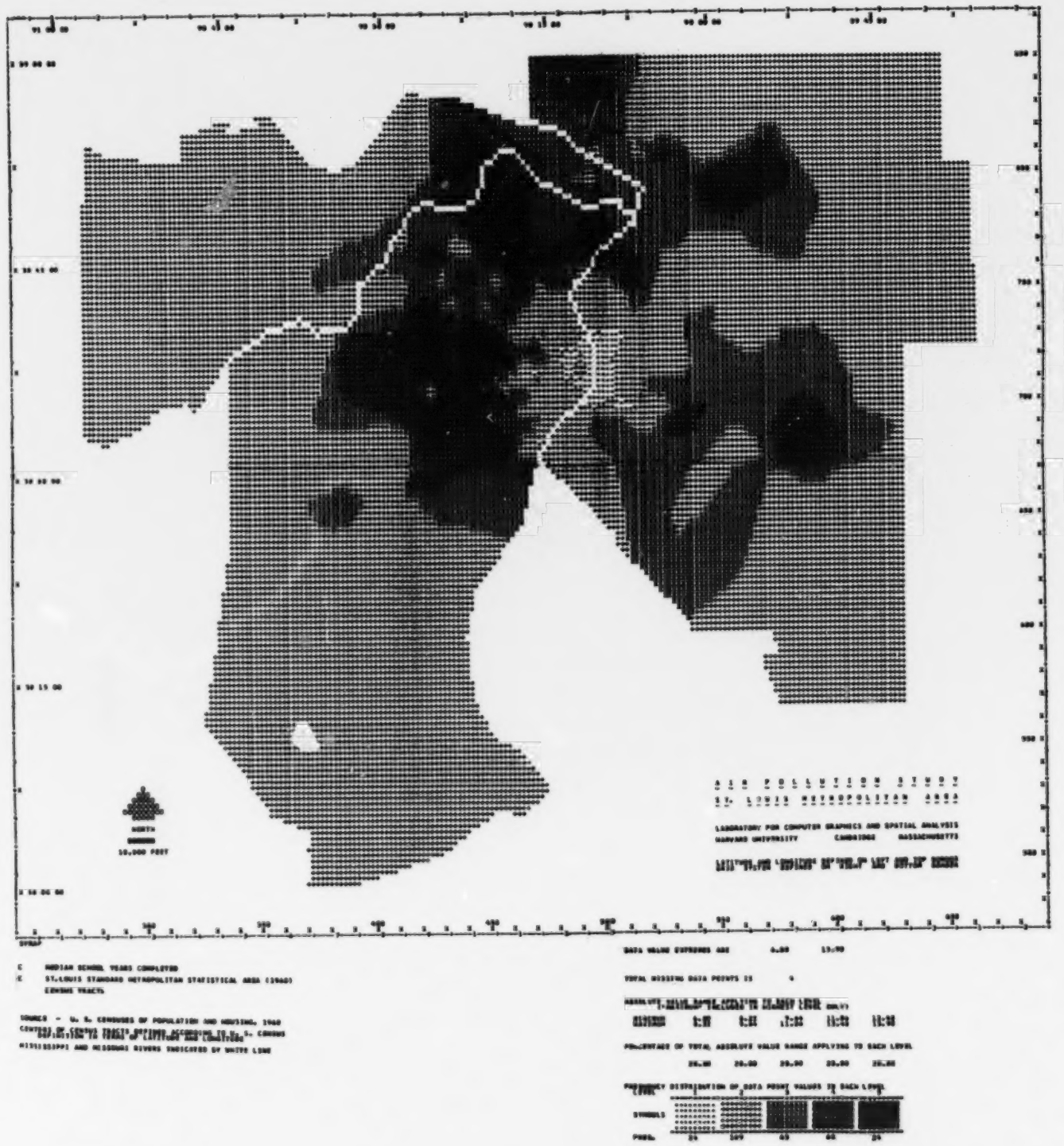


FIGURE B-13

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Median School Years Completed

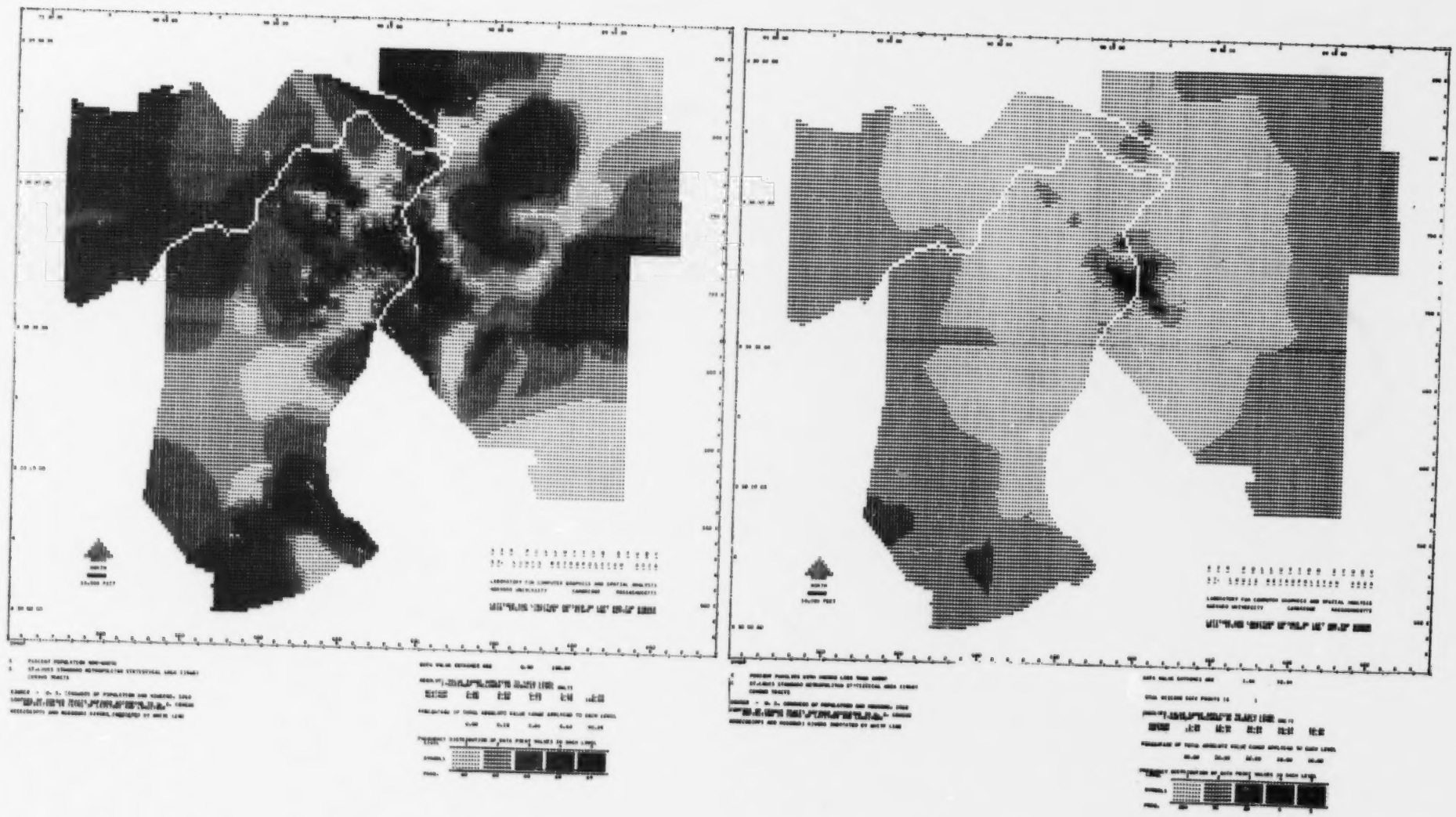


FIGURE B-14

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Percent of Population Non-white
b. Percent of Families with Incomes less than $3000/yr.

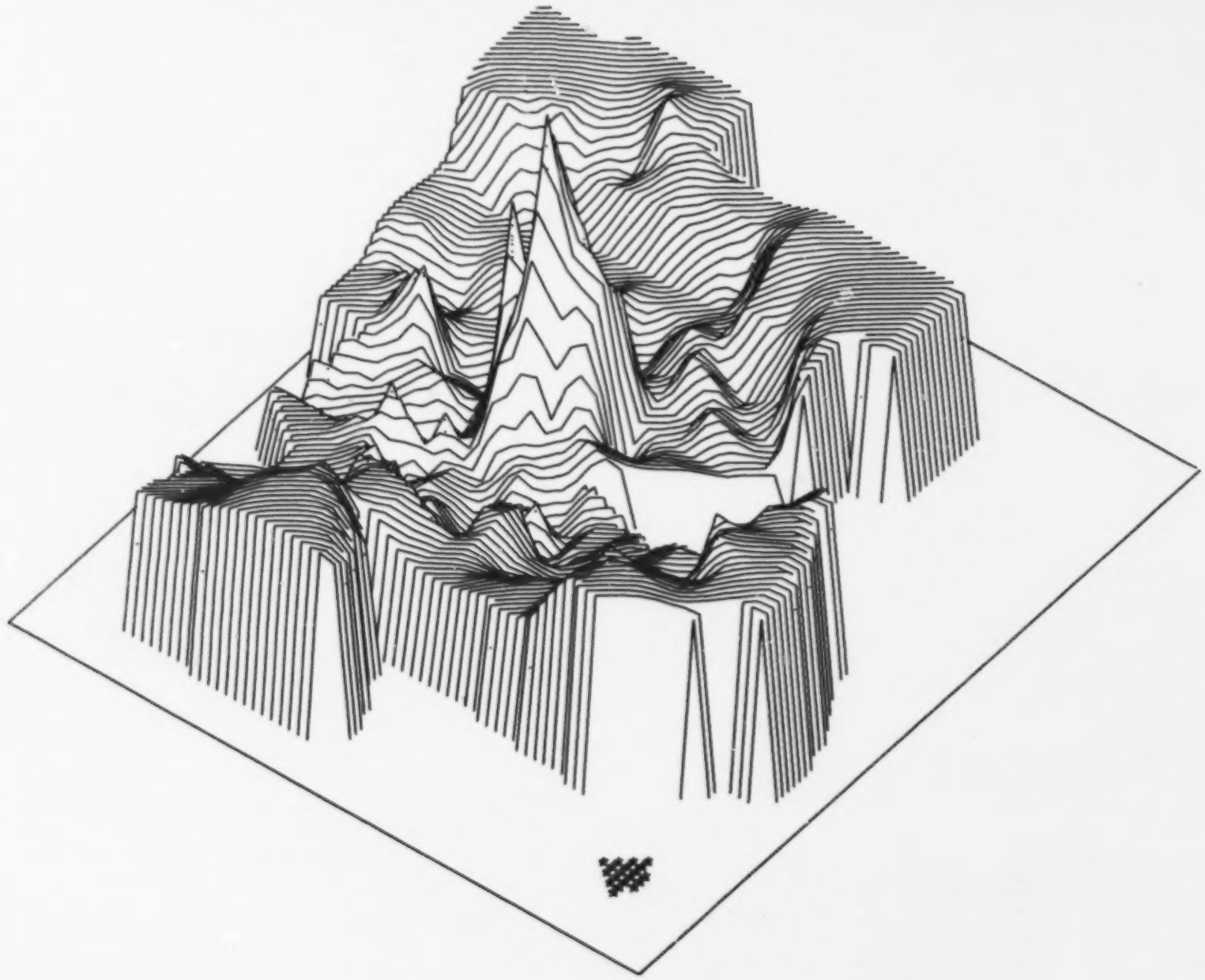
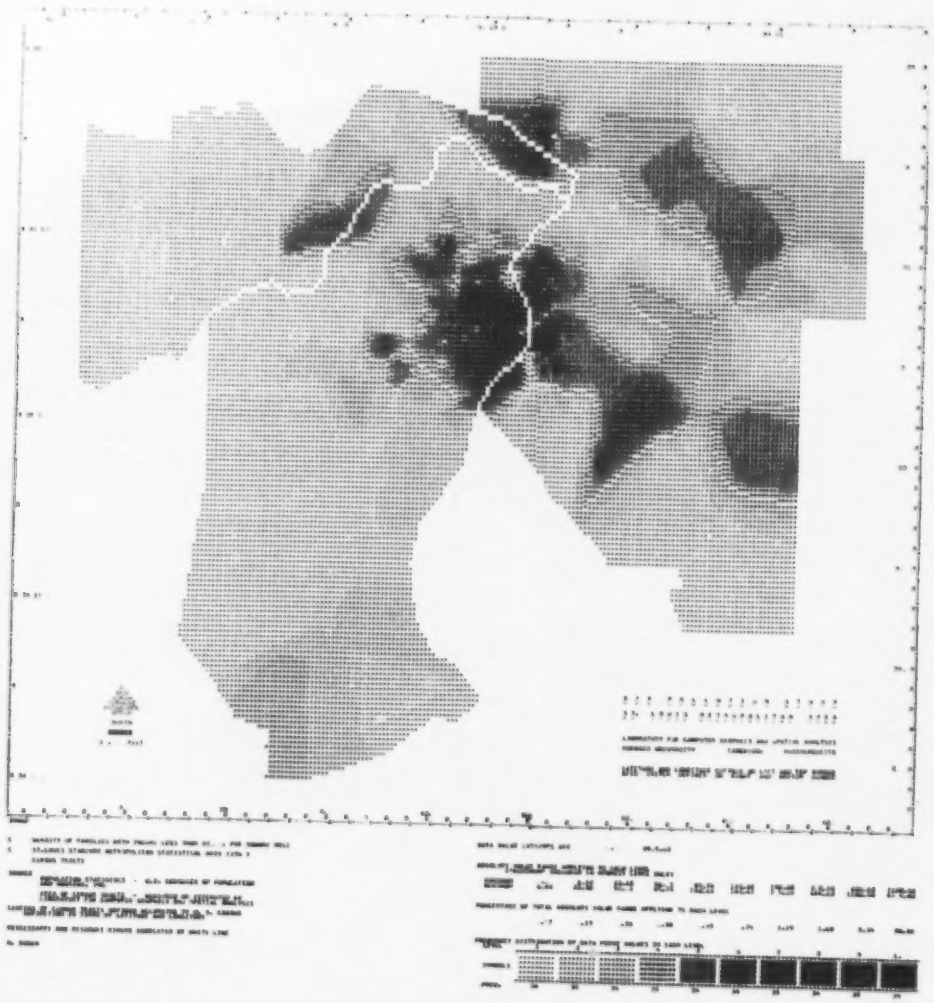


FIGURE B-15

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Families with Incomes less than $3000/yr.
Oblique View



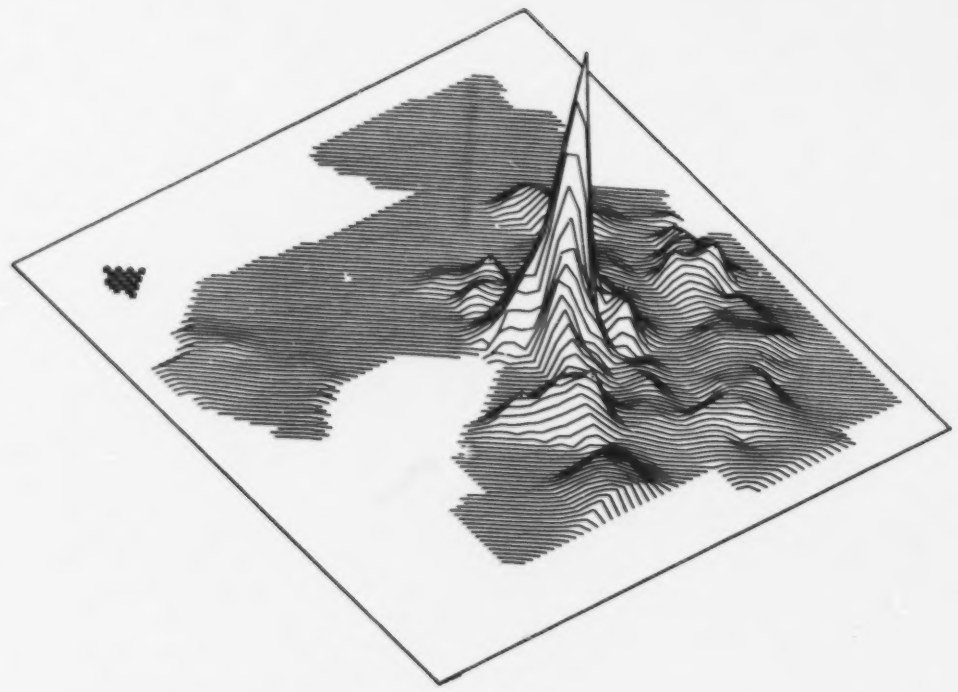


FIGURE B-16

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Density of Families with Incomes less than $3000/yr.

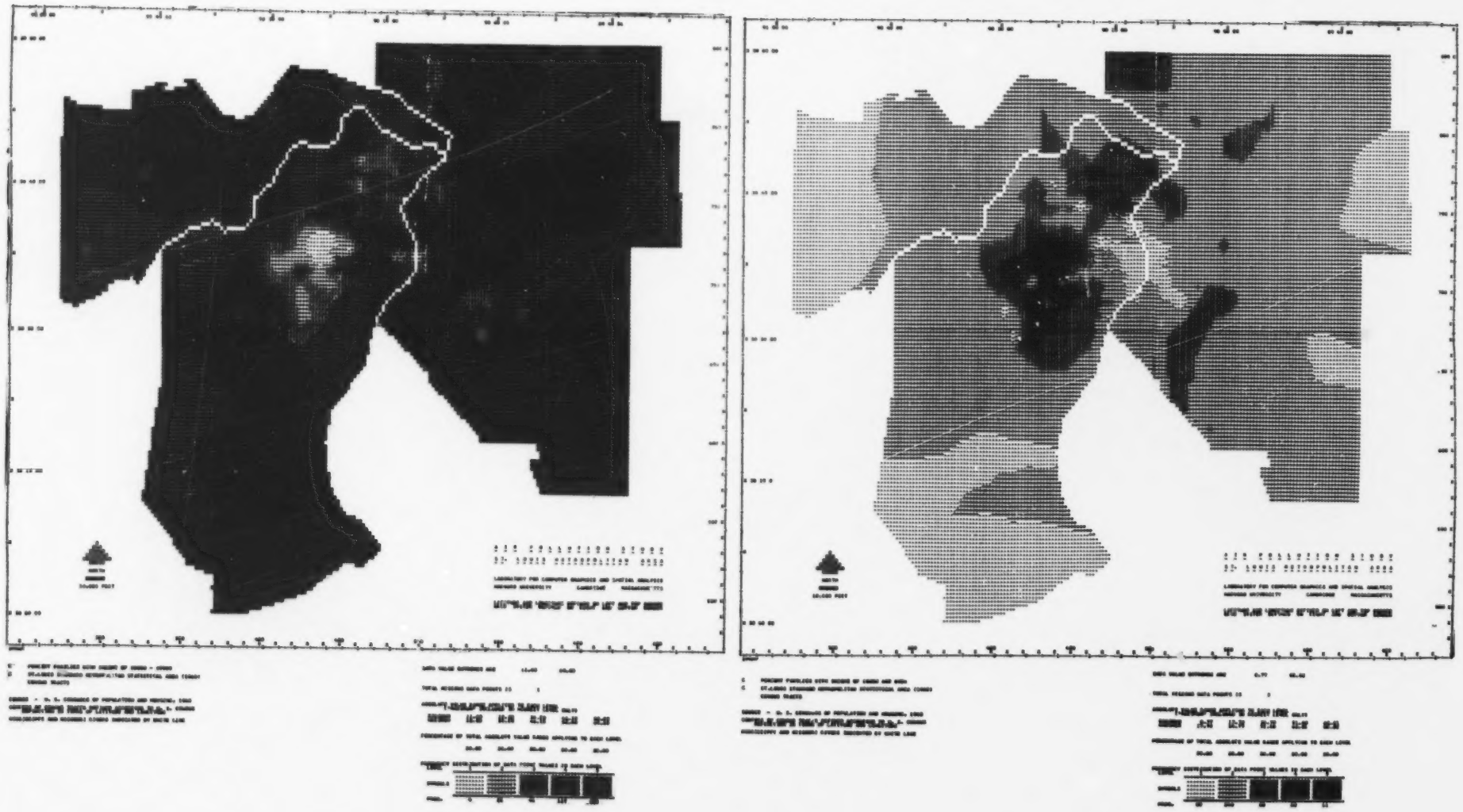
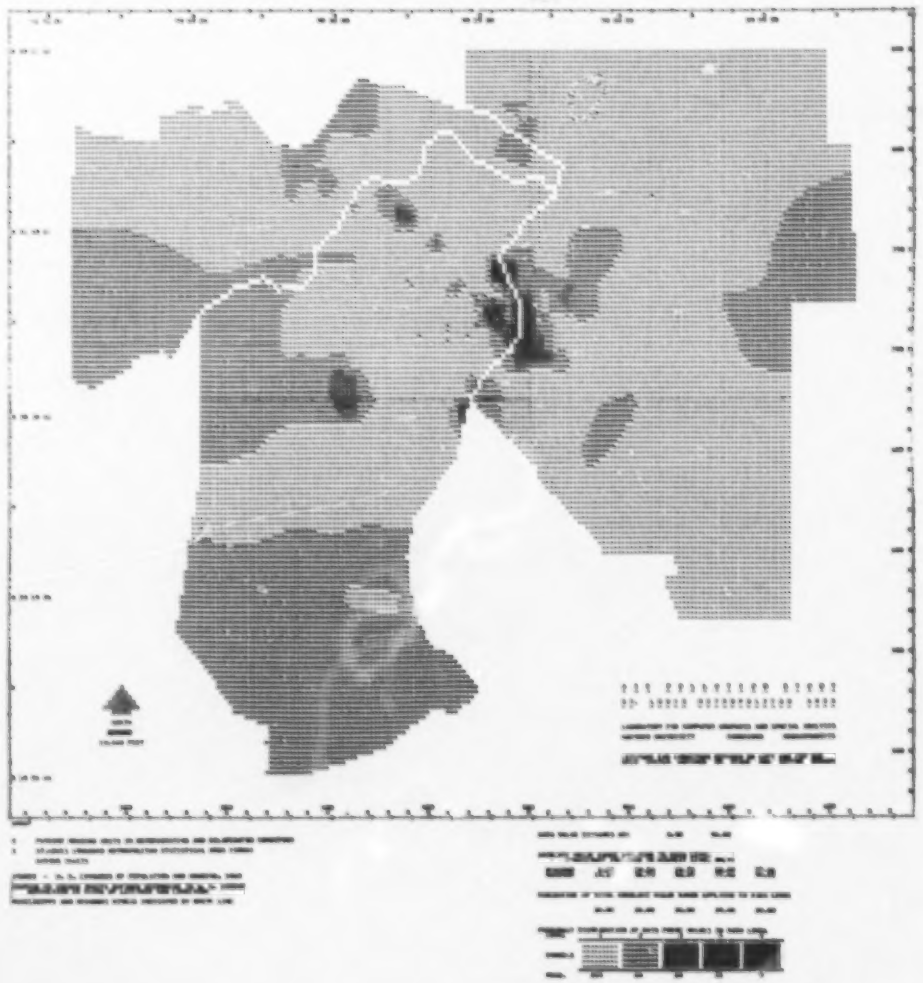


FIGURE B-17

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Percent of Families with Incomes from $3000 to $7999/yr.
b. Percent of Families with Incomes of $8000/yr. or More



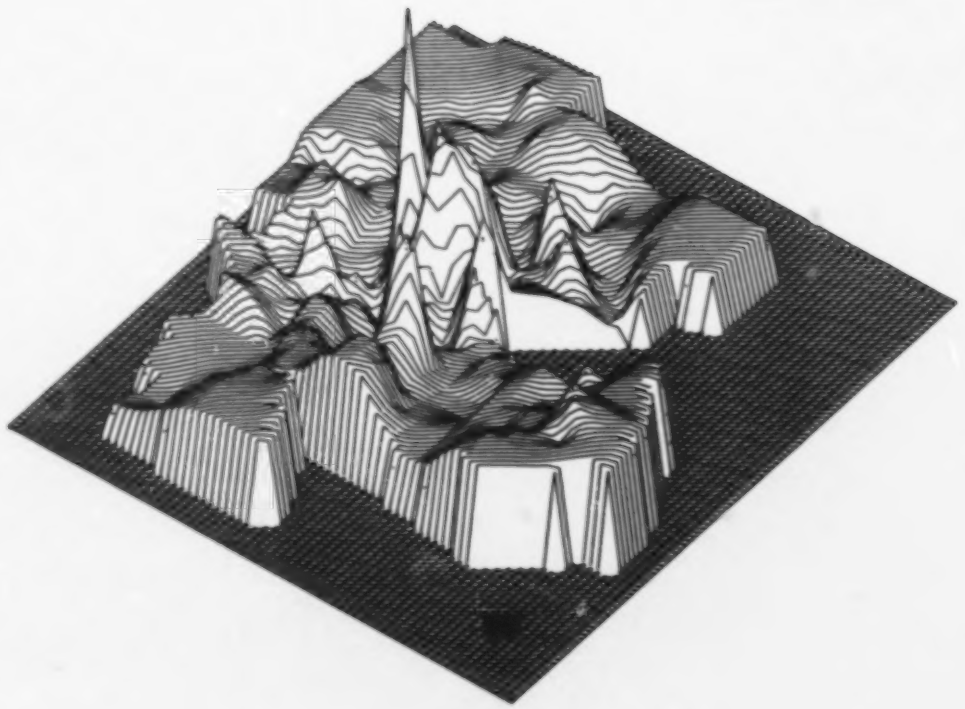


FIGURE B-18

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing Units Deteriorating and/or Dilapidated
a. Contour View b. Oblique View

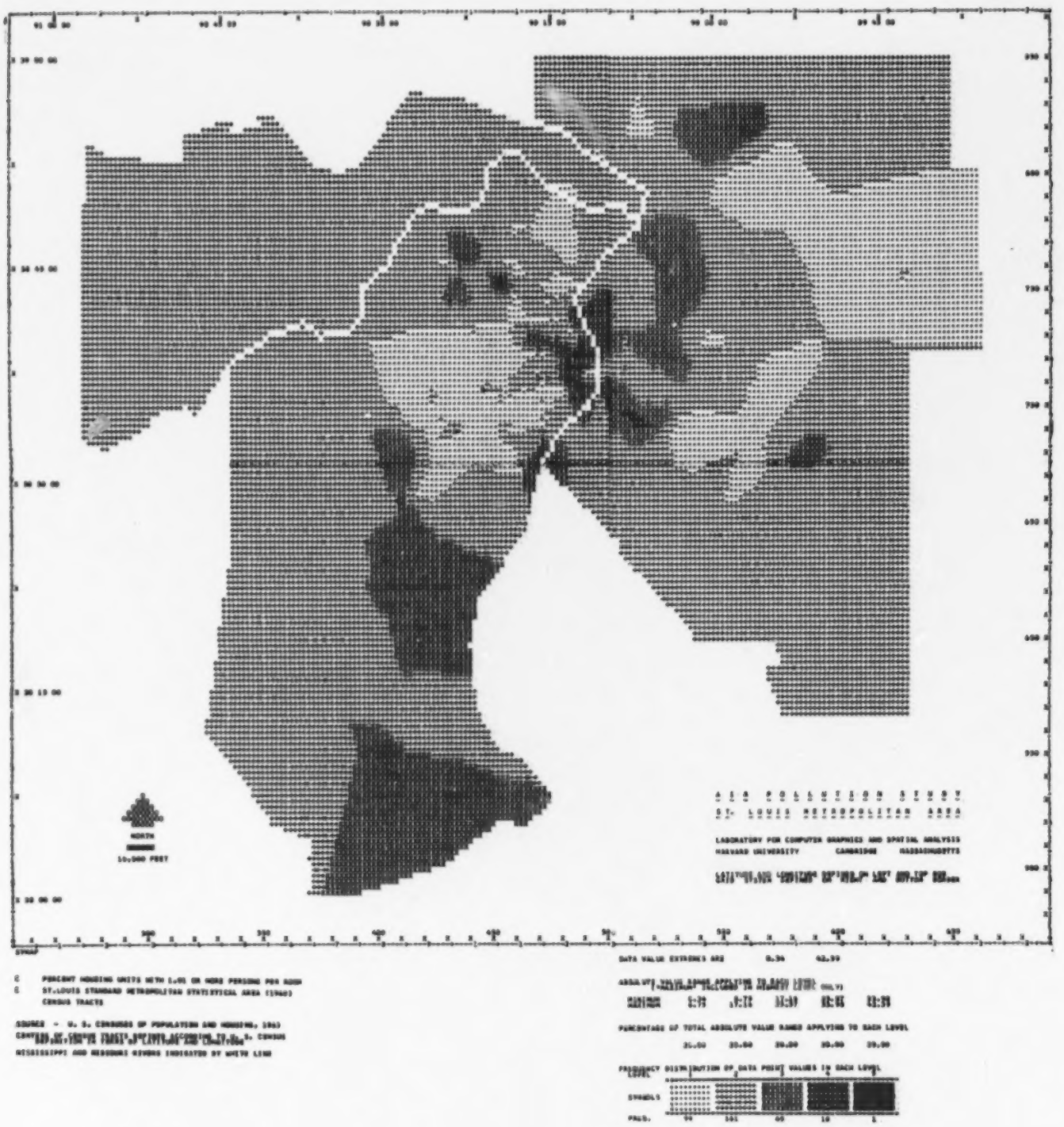

AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS
HARVARD UNIVERSITY CAMBRIDGE MASSACHUSETTS
NORTH
10,000 FEET
DATA VALUE EXTREMES ARE 0.36 62.39
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
20.00 20.00 20.00 20.00 20.00
SYMBOLS
PNTS. 56 101 65 10 1

FIGURE B-19

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing Units with 1.01 or more Persons per Room

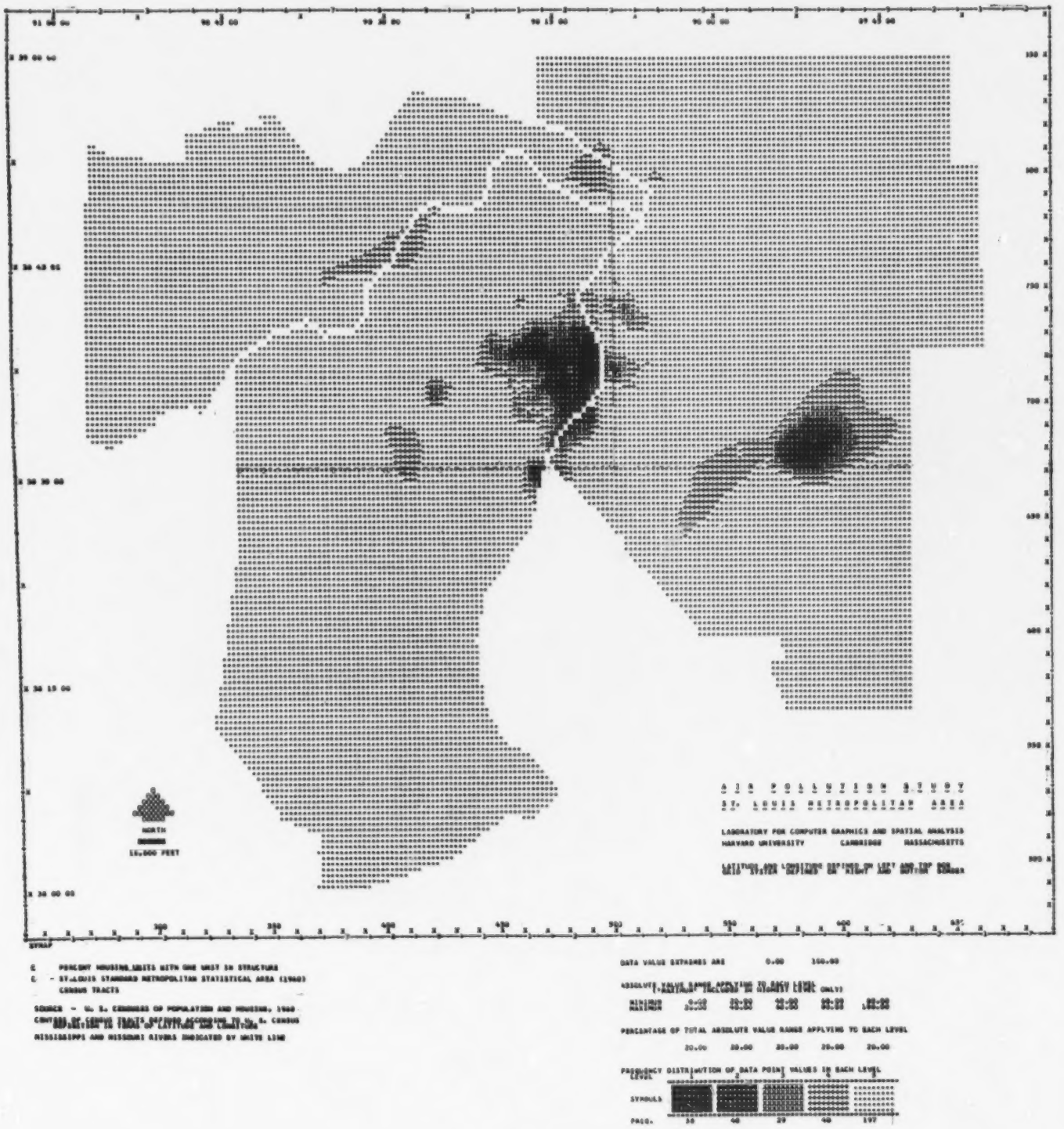


FIGURE B-20

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing with 1 Unit/Structure

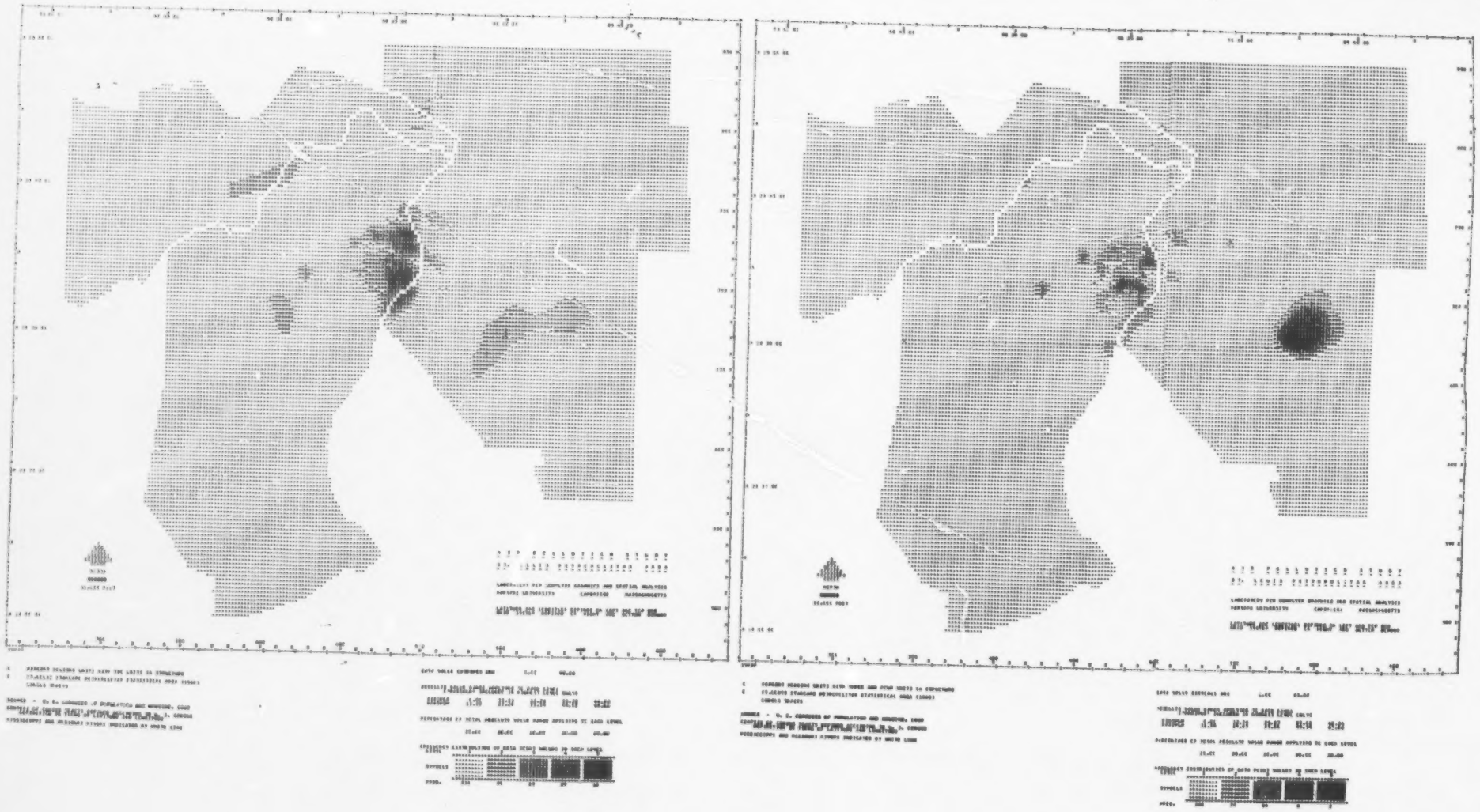


FIGURE B-21

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Percent of Housing with 2 Units/Structure
b. Percent of Housing with 3 or 4 Units/Structure

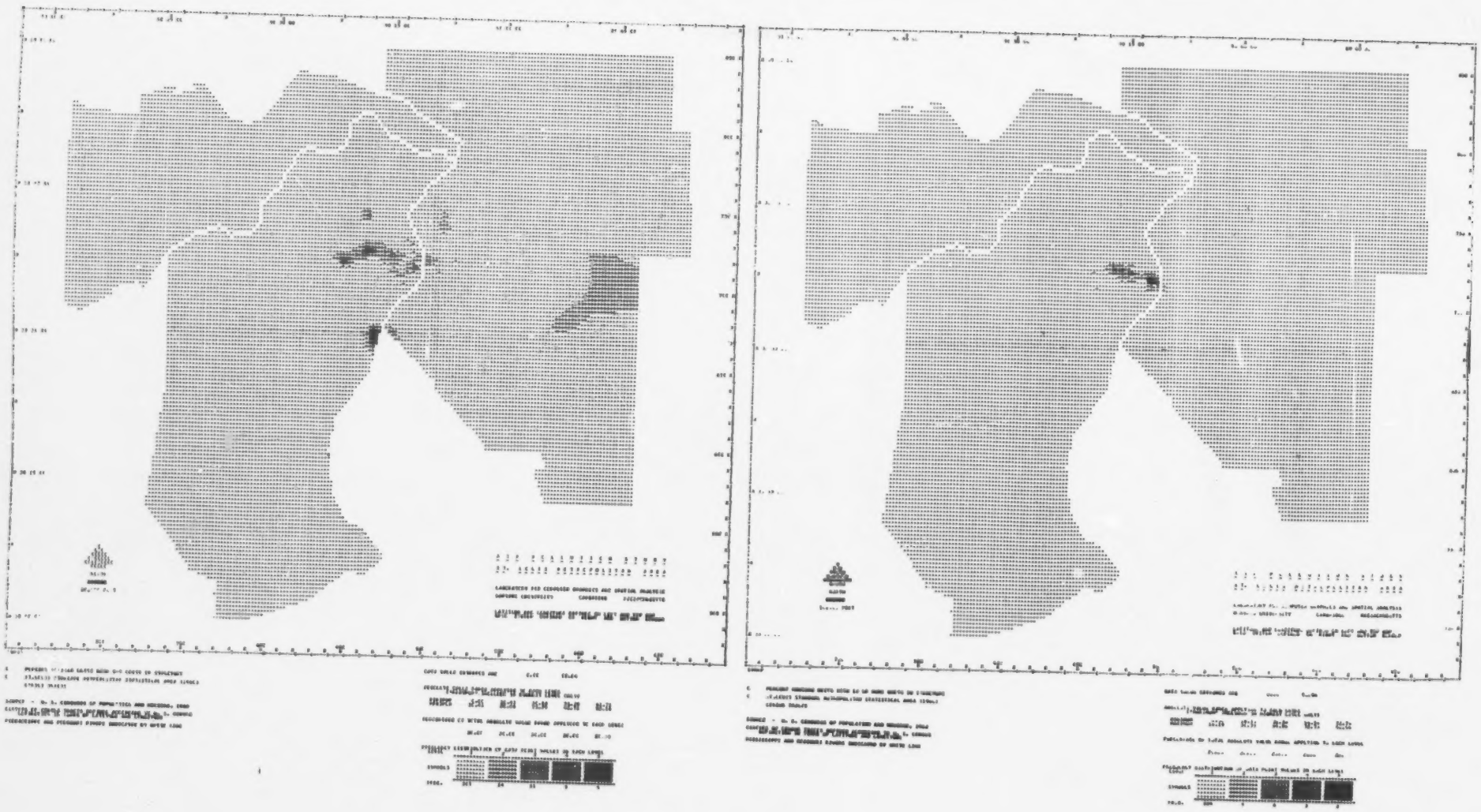


FIGURE B-22

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Percent of Housing with 5 to 9 Units/Structure
b. Percent of Housing with 10 or more Units/Structure

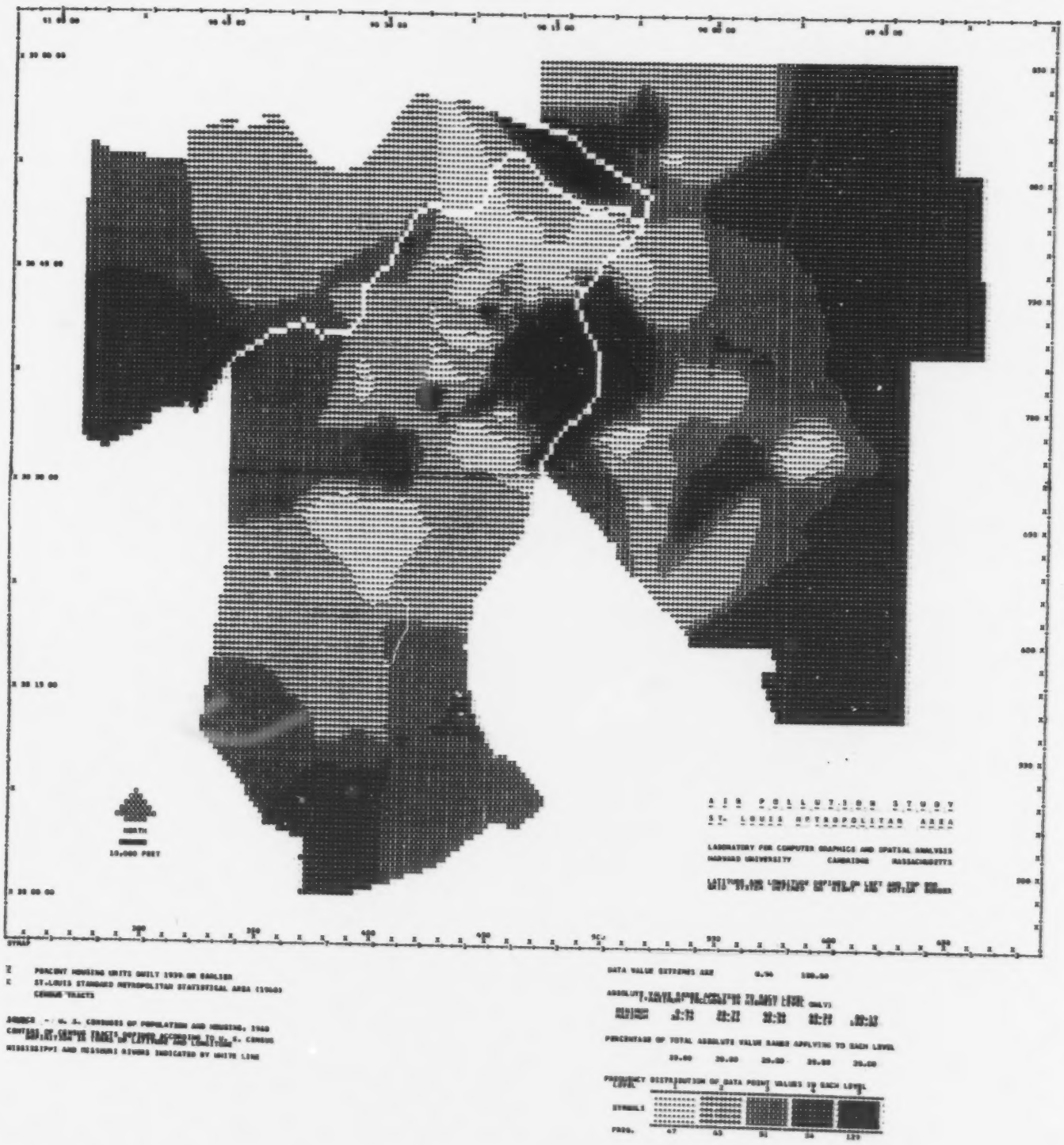
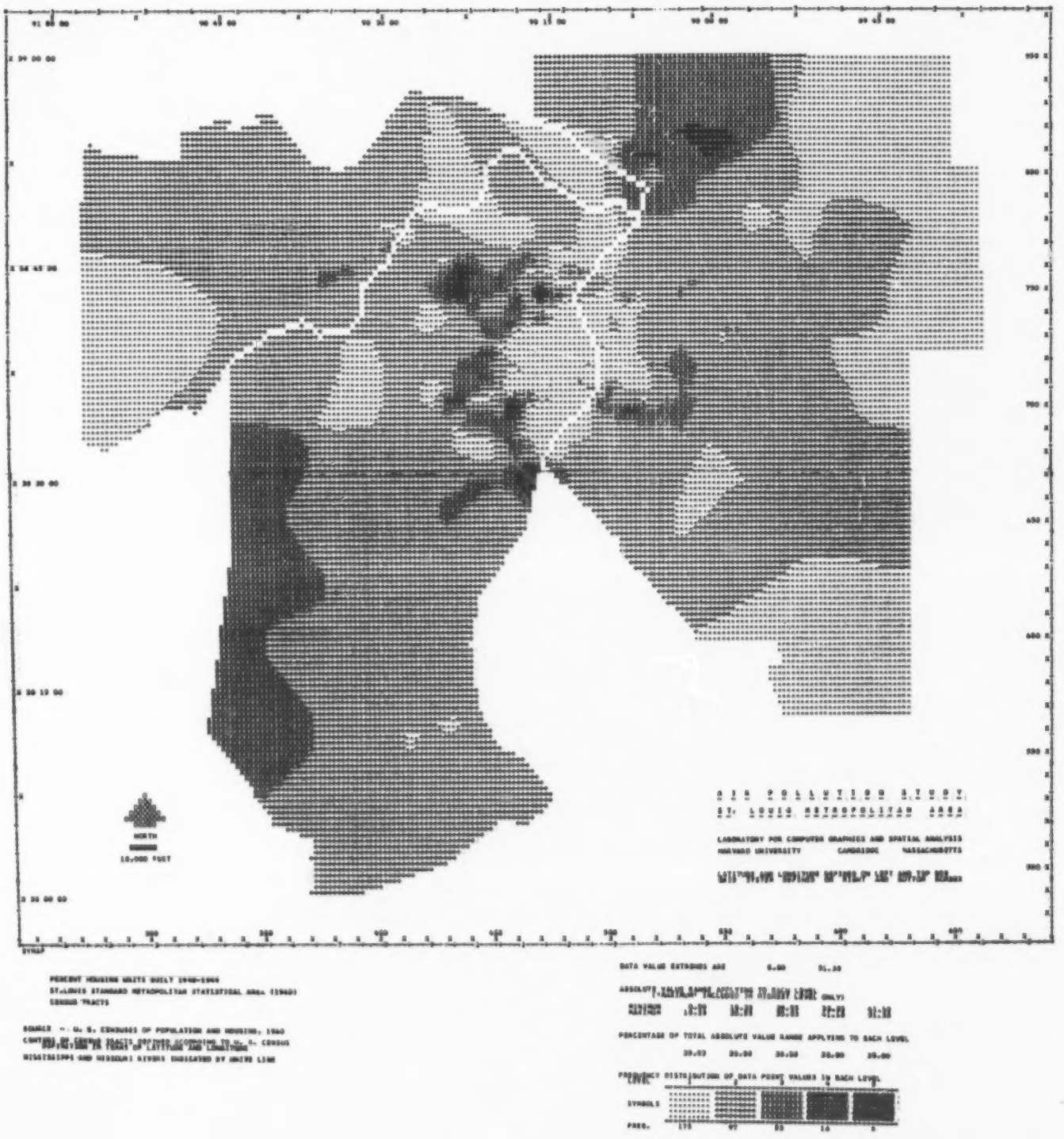

AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS
HARVARD UNIVERSITY CAMBRIDGE MASSACHUSETTS
NORTH
10,000 FEET
PERCENT HOUSING UNITS BUILT 1939 OR EARLIER
ST.LOUIS STANDARD METROPOLITAN STATISTICAL AREA (1960)
CENSUS TRACTS
DATA VALUE EXTREMES ARE 0.96 100.00
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
20.00 20.00 20.00 20.00 20.00
FREQUENCY DISTRIBUTION OF DATA POINT VALUES TO EACH LEVEL
LEVEL
SYMBOLS
FREQ.
67 63 51 24 129

FIGURE B-23

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing Built in 1939 or Earlier


AIR POLLUTION STUDY
ST. LOUIS METROPOLITAN AREA
LABORATORY FOR COMPUTER GRAPHICS AND SPATIAL ANALYSIS
HARVARD UNIVERSITY CAMBRIDGE MASSACHUSETTS
NORTH
10,000 FEET
DATA VALUE EXTREMES ARE 0.00 51.30
PERCENTAGE OF TOTAL ABSOLUTE VALUE RANGE APPLYING TO EACH LEVEL
20.00 20.00 20.00 20.00 20.00
SYMBOLS
FREQ.

PERCENT HOUSING UNITS BUILT 1940-1949
ST.LOUIS STANDARD METROPOLITAN STATISTICAL Area (1960)
CENSUS TRACTS

SOURCE - U. S. CENSUSES OF POPULATION AND HOUSING, 1960
CONVERSION OF TRACT BOUNDARIES ACCORDING TO U.S. CENSUS
MISSISSIPPI AND MISSOURI RIVERS DELINEATED BY HAVIN LINE

FIGURE B-24

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing Built 1940-1949

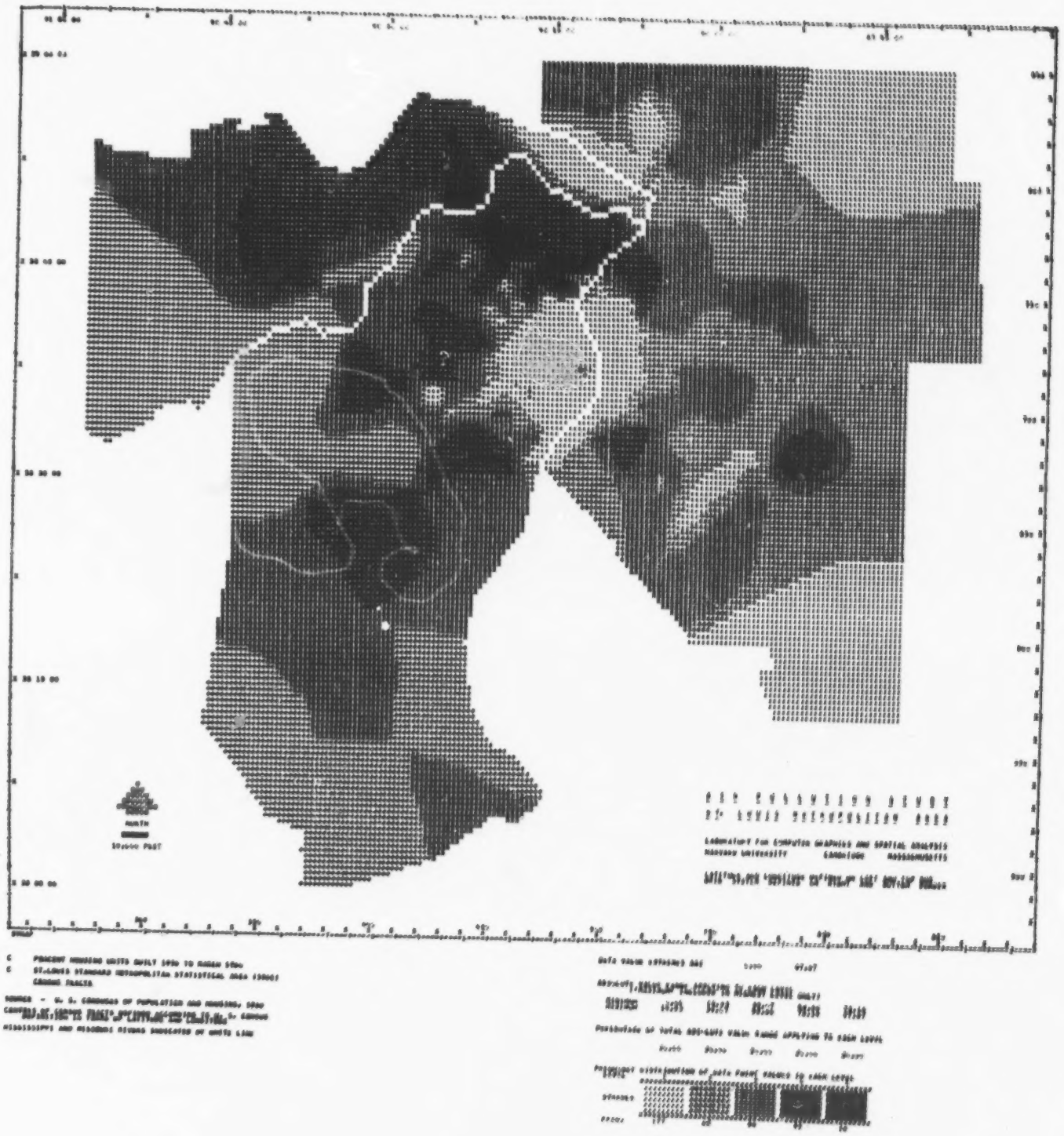
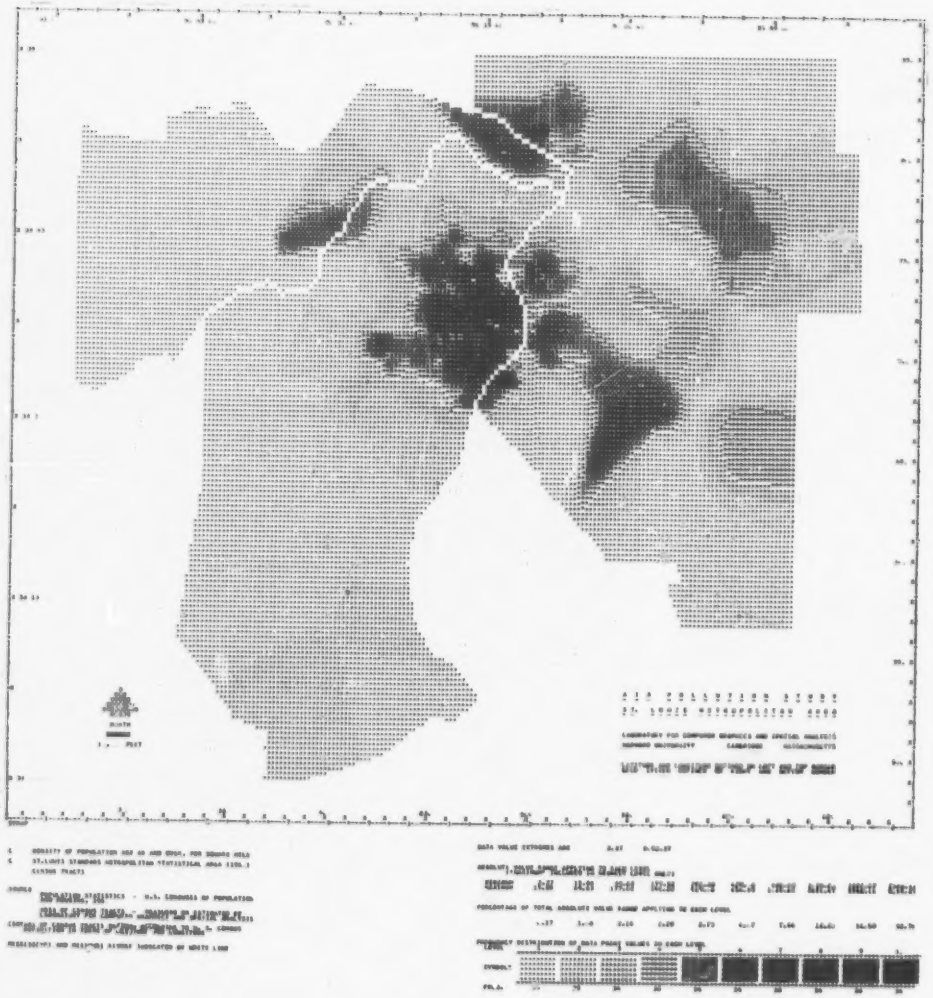


FIGURE B-25

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
Percent of Housing Built 1950-1960



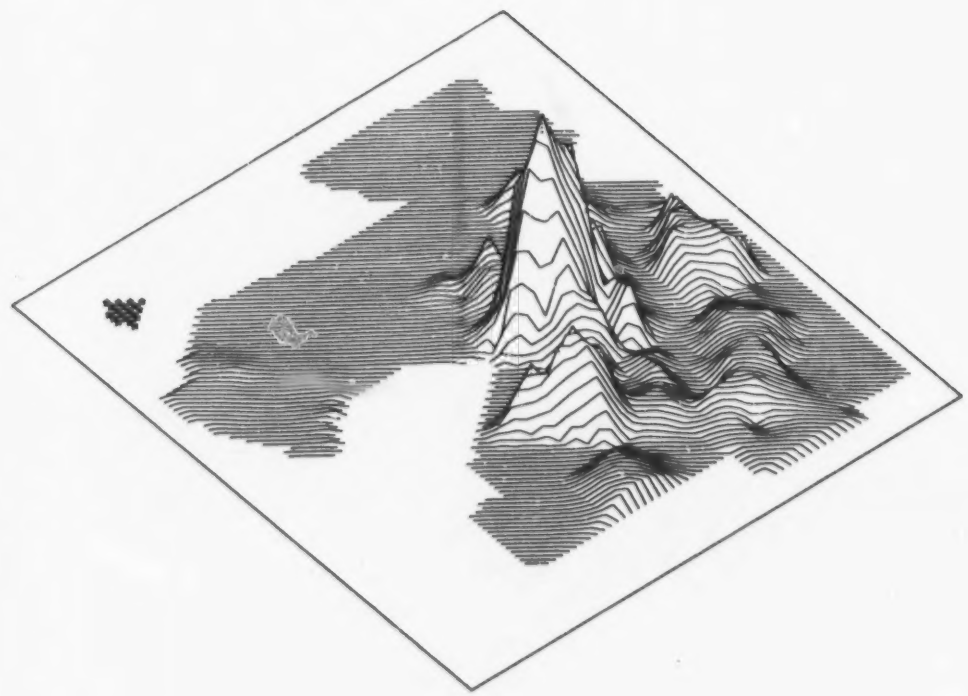
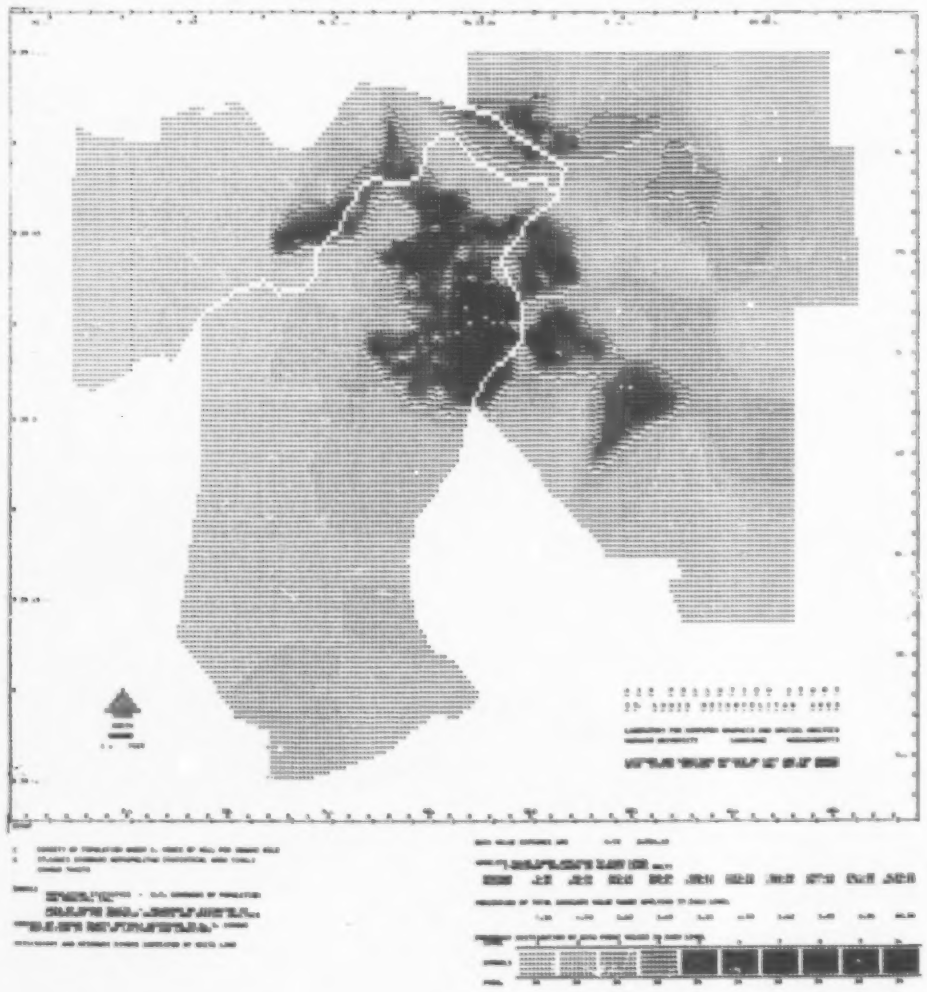


FIGURE B-26

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Density of Population over 65 Years of Age
b. Oblique View of Same



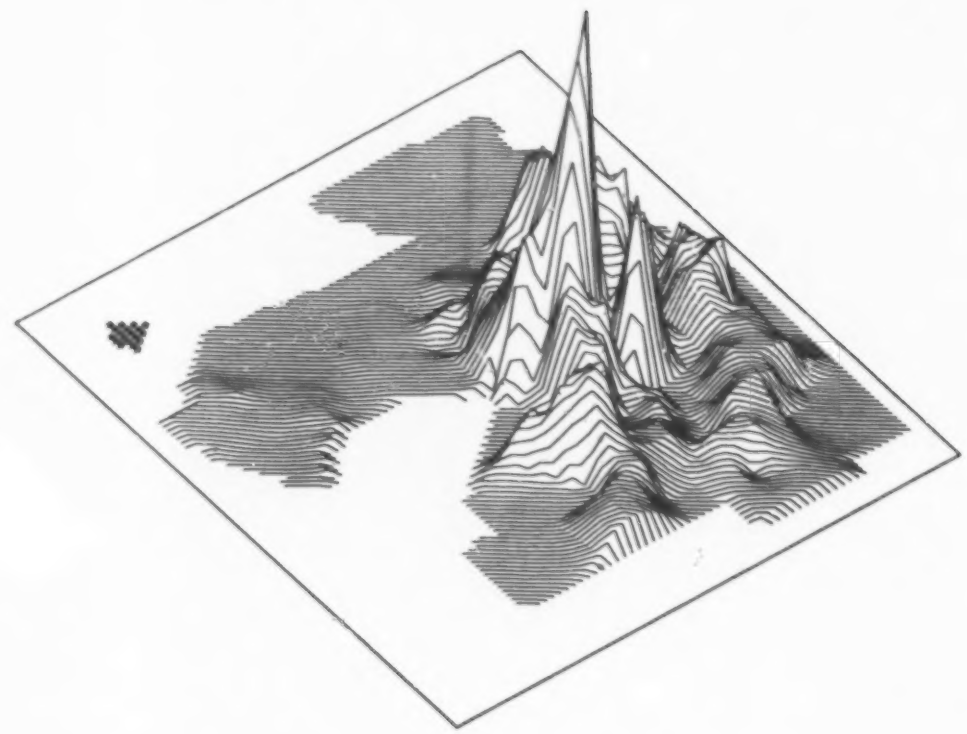


FIGURE B-27

ST. LOUIS REGION STUDY
Standard Metropolitan Statistical Area — 1960 Census
a. Density of Population under 10 Years of Age
b. Oblique View of Same

APPENDIX 2

DISCUSSION OF AIR QUALITY SURFACES

(i) Settleable particulates (dustfall) data were gathered on a monthly basis from February 12, 1963 to March 2, 1969 at forty-five sampling stations. Ideally, the data should have been collected on the same days at each of the sampling stations; however, this was a rarity. For some months there are readings at some stations and none at others. For comparability the samples should have been of equal or nearly equal duration, but there were samples taken over much less and much more than a month's duration.

(ii) Suspended particulates (hi-vol) data were gathered from May 20, 1963 to April 28, 1964 at seventeen sampling stations. One twenty-four hour reading should have been taken each day. It appears that many of the readings were taken for approximately one or two hours rather than for a twenty-four hour period. Sampling at most of the stations was done at about the same time each day. Where this was not the case, it is difficult to compare the data.

(iii) Soiling index (ASI) readings were taken from May 20, 1963 to April 21, 1964 at twelve sampling stations. At some stations the readings covered several months, while at others only a few months were included. All stations had missing values -- two-hour periods for which no reading was recorded.

(iv) Sulfation was recorded from February 12, 1963 to March 2, 1964 at fifty sampling stations. Sulfation readings should be taken once a month, but some of the readings encompassed

less than a month and others more than a month. In addition, some stations have records for only a few months. An attempt was made to take readings on a particular day of a month at all stations.

(v) Sulfur dioxide two-hour sequential readings were taken from September 4, 1963 to February 26, 1965 at eleven stations. This appears to be one of the most carefully collected data sets, although there are missing readings.

(vi) Twenty-four hour sulfur dioxide readings were also recorded, but only during the winter of 1965 -- from November 30, 1964 to February 27, 1965. Forty sampling stations were used and the observations were made at the same time each day. There are several days for which no data were recorded.

(vii) Oxident (phenolphthalin) data were gathered from May 21, 1964 to October 14, 1964 at eight stations. One of the stations has very incomplete records.

The larger the number of sampling stations at which observations on a particular pollutant are made, the better the final estimates will be. Unfortunately, the pollutants which were measured on a large scale were those that could be measured cheaply and not those which yield the most useful information. For example, sulfation was measured at the greatest number of stations, but was one of the least interesting variables. Most of our analysis was done with the sulfur dioxide data -- measured at only a few points -- because sulfur dioxide is considered one of the most significant pollutants. Naturally, the scarcity of data affects the analysis.

APPENDIX 3

DISCUSSION OF SOCIO-ECONOMIC SURFACES

(i) Population density was highest in the central city, decreasing in all directions; several smaller peaks appeared in the outer part of the region. (Figures B-9 and B-10)

(ii) The pattern for the percent of population under ten years old was found to be high only in very small portions of the central city, and higher in the suburbs. (Figure B-11)

(iii) The percent of population over sixty-five years old was highest in the central city, quite low in a large suburban ring around the central city, and high again in the rural areas. (Figure B-12)

(iv) Median school years completed was one of those variables that was included on the chance that it might prove to be of some use. The major peaks of this surface occur west of the central city. The variable was very highly correlated -- .859 -- with income over $8000 as one might expect, but it was not very useful for comparing with the air pollution surfaces. (Figure B-13)

(v) The percent of the population non-White proved to be quite a significant variable. It behaved very differently from the other population variables: in general, it was either significant in a region or almost non-existent. There was an extended peak in the central city and a couple of very small suburban peaks, with nearly all other areas having less than twenty percent non-White. (Figure B-14a)

(vi) Percent of the population having incomes under $3000 was another variable with a high central peak, lower values in the suburbs, and high values on the outer fringe of the area. (Figures B-14b and B-15) Compare these maps with the corresponding density maps (Figure B-16)

(vii) Percent of the population having incomes between $3000 and $7999 is one of the more difficult surfaces to describe. It was high in the suburbs except for the highest income area. It also peaked in the central city and on the fringes. (Figure B-17a)

(viii) Percent of the population having incomes of $8000 and over peaked west of the central city. (Figure B-17b)

(ix) Percent of housing deteriorated or delapidated was highly correlated (.889) with the location of low income families, as would be expected. It was highest in the central city and had some scattered peaks in the suburbs and on the fringes. (Figure B-18)

(x) Percent of housing with 1.01 or more persons per room is a measure of crowding. This had a high peak in the central city and a large low area where the peak of the surface of high incomes occurred. The rest of the surface was composed of many local peaks and troughs. (Figure B-19)

(xi) Percent of housing with one unit in the structure is most indicative of suburban living. There was a large low area in the central city surrounded by high levels in the suburbs. (Figure B-20)

(xii) Percent of housing with two units in the structure had a peak east of the city and two other minor peaks. In the major subsidiary peak east of the city the percent of one unit dwellings was low. (Figure B-21a)

(xiii) Percent of housing with three or four units per structure greatly resembled the surface for two unit structures. (Figure B-21b)

(xiv) Percent of housing with five to nine units per structure produced a surface with localized peaks in the central city. Its major characteristic was a band running east to west in the northern part of the central city, starting at the river. There were also some small peaks to the north and south. (Figure B-22a)

(xv) Percent of housing with ten or more units per structure looked like the surface for five to nine units. It had only the band, however, and no subsidiary peaks. (Figure B-22b)

(xvi) The percent of housing built in 1939 or earlier was high in the central city, fairly low in the suburbs, and high again in the rural areas. (Figure B-23)

(xvii) The percent of housing built from 1940 to 1949 was high in many suburban areas. (Figure B-24)

(xviii) The percent of housing built from 1950 to 1960 was high in the suburbs and low in the central city, except for a very small peak near the center. (Figure B-25)

Two additional population variables were used in conjunction with the particulates data:

(xix) The density of the population over sixty-five was highest in the central city, with some peaks in surrounding areas. (Figure B-26)

(xx) The density of population under ten years of age resembled the above surface in general. (Figure B-27)